UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION: Unclassified

1b. RESTRICTIVE MARKINGS:

2a. SECU

2b. DEC

4. PERF

AD-A228 147

3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release; distribution unlimited.

5. MONITORING ORGANIZATION REPORT NUMBER(S): ARO 24543.13-EL

6a. NAME OF PERFORMING ORGANIZATION:

6b. OFFICE SYMBOL *(If applicable)*:

7a. NAME OF MONITORING ORGANIZATION: U. S. Army Research Office

6c. ADDRESS *(City, State, and ZIP Code)*:

7b. ADDRESS *(City, State, and ZIP Code)*:
P. O. Box 12211
Research Triangle Park, NC 27709-2211

8a. NAME OF FUNDING/SPONSORING ORGANIZATION: U. S. Army Research Office

8b. OFFICE SYMBOL *(If applicable)*:

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: DAAL03-87-K-0062

8c. ADDRESS *(City, State, and ZIP Code)*:
P. O. Box 12211
Research Triangle Park, NC 27709-2211

10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.

11. TITLE *(Include Security Classification)*:
Near/Far Resistant Receivers for DS/SSMA Communcations (Final Report)

12. PERSONAL AUTHOR(S): Sergio Verdú

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
FINAL	FROM 5/1/87 TO 7/31/90	1990.9.27	

16. SUPPLEMENTARY NOTATION: The view, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

17. COSATI CODES

FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS *(Continue on reverse if necessary and identify by block number)*:
Spread Spectrum, Direct-Sequence, Near-Far Problem
Detection Theory, Multiuser Communication

19. ABSTRACT *(Continue on reverse if necessary and identify by block number)*

Final report including a brief summary of main research findings, a list of publications sponsored under this contract and a list of scientific personnel involved in the project.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: ☐ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS

21. ABSTRACT SECURITY CLASSIFICATION: Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL:

22b. TELEPHONE *(Include Area Code)*:

22c. OFFICE SYMBOL:

U. S. ARMY RESEARCH OFFICE

FINAL REPORT

for

Contract DAAL03-87-K-0062

1 April 1987 --- 31 July 1990

NEAR/FAR RESISTANT RECEIVERS FOR DS/SSMA COMMUNICATIONS

Sergio Verdú, Principal Investigator
Department of Electrical Engineering
Princeton University
Princeton, New Jersey 08544
Phone: (609) 258-5315
Email: verdu@princeton.edu

TABLE OF CONTENTS

Summary of Main Results

In Code-Division Multiple-Access (CDMA), each transmitter is assigned a fixed, distinct signature waveform which he uses to modulate his message in the same fashion as in single-user communication. Then the information sent by each user can be demodulated by correlating the received signal with each of the signature waveforms. This demodulator, whose use is widespread in practice, is referred to as the conventional single-user detector. As is well-known, when the channel output is corrupted by additive white Gaussian noise, the conventional single-user detector minimizes the probability of error in a single-user channel, i.e., in the absence of interfering users. The fact that this is no longer true in the multiple-access channel is the raison d'être of the area of multiuser detection.

The performance of the conventional single-user detector is acceptable provided that the energies of the received signals are not too dissimilar and that the signature waveforms are designed so that their crosscorrelations are low enough (this depends on the desired maximum number of simultaneous users). In practice, low crosscorrelations are usually achieved employing Spread-Spectrum Pseudonoise sequences of long periodicity. If the received signal energies are indeed dissimilar, i.e., some users are very weak in comparison to others, then the conventional single-user detector is unable to recover the messages of the weak users reliably, even if the signature waveforms have very low crosscorrelations. This is known as the *near-far problem* and is the main shortcoming of currently operational Direct-Sequence Spread-Spectrum Multiple-Access systems, and of recently proposed systems for future mobile radio communications.

Due to the reduction of multiple-access capability and the increase of vulnerability to hostile sources caused by the near-far problem, its solution or alleviation had been a

target of researchers in the area for several years. Before the emergence of the solutions based on multiuser detection developed by the Principal Investigator and his coworkers, success had been very limited and, essentially, the only remedies available were power control and the design of signals with even more stringent crosscorrelation properties. Unfortunately, power control (i.e., the adaptive adjustment of transmitter power depending on its location and of the received powers of the other users) dictates significant reductions in the transmitted powers of the strong users in order for the weaker users to achieve reliable communication. Thus, power control can become self-defeating since it actually decreases the overall multiple-access and antijamming capabilities of the system. Furthermore, more and more complex signature waveforms lead to rapid increases in system cost and bandwidth, and, as we have noted, do not eliminate the near-far problem. For these reasons, it can be seen why the practical solution to the near-far problem achieved in this project can be objectively considered a major breakthrough in the application of signal processing techniques to Spread Spectrum communications.

The chief reason why multiuser detection did not develop until relatively recently was the belief shared by many a worker in Spread-Spectrum that multiuser interference is accurately modeled as a white Gaussian random process, and thus the conventional detector is essentially optimum. It is not difficult to build an infinite population multiuser signal model which can be rigorously shown to be asymptotically Gaussian as the individual amplitudes go to zero with the appropriate speed. Unfortunately, the number of transmitters, signature waveforms, and power levels encountered in many practical situations (e.g. in near-far environments) render the Gaussian approximation completely useless. Therefore, it is useful to adopt a more refined viewpoint by taking the realities of the medium into account, modeling them and exploiting them.

Prior to the start of this project, it had been shown by the Principal Investigator in

"Minimum Probability of Error for Asynchronous Gaussian Multiple Access Channels," *IEEE Trans. Information Theory,* Jan. 1986 and "Optimum Multiuser Asymptotic Efficien y," *IEEE Trans. Communications,* Sep. 1986, that the near-far problem is not an inherent flaw of Direct-Sequence Spread-Spectrum systems but rather of the simple conventional correlation receiver. Those works developed optimum multiuser receivers, which consist of a front end of matched filters followed by a Viterbi-type algorithm whose number of states is exponential in the number of users. Those receivers did not suffer from the near-far problem. However, their implementation suffered from the burden of the exponential complexity in the number of interferers and from the necessity to estimate the received powers of each interferer.

The first breakthrough obtained in this project was the demonstration that for *symbol-synchronous* multiple-access channels a simple modification of the conventional correlation receiver results in a system which does not exhibit the near-far problem, and moreover provides optimum robustness against variations in the received strength of the various transmissions. We called this new receiver the *decorrelating detector* as it correlates the received signal against a linear combination of the waveforms assigned to the active users, rather than only the waveform of the user of interest. The coefficients of such a linear combination do not depend on the relative strengths of the transmitters and can be precomputed in advance. Such linear combination is such that the detector correlates with the projection of the signature waveform of the user of interest on the subspace orthogonal to the subspace spanned by the interfering waveforms. Thus, the decorrelating detector effectively tunes out the multiuser interference.

A very pleasant surprise was to demonstrate that the decorrelating detector achieves optimum near-far resistance, i.e. the same level of protection against the near-far problem as the optimum detector. This means that knowledge of the received energies is not required to combat the near-far problem and that a receiver whose complexity is similar to that of the conventional detector achieves the same degree of robustness against imbalances in the received energies as the optimum detector, with its exponential complexity. Another attractive property of the decorrelating detector is that its bit-error-rate is independent of the energies of the interfering transmitters--a most desirable feature of a strategy designed to combat the near-far problem. The signal-to-noise reduction due to the presence of interfering waveforms depends exclusively on the (crosscorrelations of the) signature waveforms assigned to the transmitters, and not on their relative power levels.

The next major result was the development of the decorrelating detector for the asynchronous Code-Division multiple-access channel. In contrast to the solution obtained for the synchronous channel, the asynchronous decorrelating detector is a linear system with memory which can be implemented by transversal discrete-time filters. Although the technical development and bit-error-rate analysis is complicated by an order of magnitude in the asynchronous case, all the desirable features of the synchronous decorrelating detector are retained: optimal near-far resistance, independence of bit-error-rate to relative power levels, low complexity, and the fact that it is unnecessary to know the received energies of the individual interferers at the receiver.

The main increase in complexity of the decorrelating detector with respect to that of the conventional matched filter is the need to lock to the signal epochs of the interfering users, which can be accomplished by a bank of conventional single-user synchronization systems, or (in a promising direction for future research) by multiuser synchronizers.

We have also developed low-complexity approximations to the decorrelating detector, which do not require the iterative or off-line computation of equalizer taps and perform promisingly for signature waveforms such as those used in currently operational

Direct-Sequence Spread-Spectrum systems.

Although the decorrelating detector does not require knowledge of the powers of the interfering transmitters, we have found a class of related linear receivers which makes effective use of this information. Consequently, we have tackled the problem of amplitude estimation in a multiuser white Gaussian channel and have suggested promising techniques that can be brought to bear on this problem. Those techniques can be categorized as being based either on maximum-likelihood data recovery or on minimum error-probability (i.e., Bayesian) data recovery. In each case, the most promising algorithm is iterative with Gauss-Seidel iteration and the EM algorithm being proposed for the ML and MEP approaches, respectively.

We have also considered nonlinear receivers for multiuser channels. Here, each receiver may only know the signature sequence of one of the transmitters and treats the sum of all the other transmitted waveforms as noise, which is neither white nor Gaussian. Particular emphasis is placed on asymptotically optimum detectors for each of the following situations: weak interferers; CDMA signature waveforms with long spreading codes; and low background Gaussian noise level.

We have also looked at issues of computational complexity of combinatorial optimization problems arising in multiuser detection. We have shown that minimum bit-error-rate demodulation of multiuser signals is an NP-complete problem in the number of users, and hence polynomial algorithms are out of the question unless well-known problems such as the traveling-salesman and integer linear programming can also be solved in polynomial time. Fortunately, however, minimum bit-error-rate receivers are not the only ones that are resistant against the near-far problem. In particular, when, in lieu of bit-error-rate, near-far resistance is the optimality criterion, then the complexity of the optimal receiver (the decorrelating detector) is linear in the number of interferers.

When Spread-Spectrum signaling is coupled to random-access protocols to provide some degree of flow control, it is necessary to do away with the conventional collision channel model, whereby several simultaneous transmissions result in the destruction of all transmitted messages. We have proposed a new model, the *multipacket channel*, which is general enough to encompass random-access channels with CDMA or with capture (which are especially relevant in near-far situations). More significantly, we have found the maximum throughput achievable by the ALOHA algorithm in the general multipacket channel: 1) in the open-loop version of the ALOHA algorithm, i.e., with fixed retransmission probabilities, the throughput is equal to the limit of the expected number of successfully received packets per slot as the backlog goes to infinity, and 2) the throughput of closed-loop ALOHA, where the retransmission probabilities are a function of the channel outcomes, is equal to the maximum over υ of the expected number of successfully received packets per slot when the number of attempted transmissions is a Poisson random variable with mean υ.

Our research efforts under the sponsorship of this ARO contract have also been directed to the investigation of related issues in the CDMA optical channel, which is receiving considerable attention both in commercial and military applications for fiber-optic and free-space photonic channels. The emphasis has been in the development of formulas for the bit-error-rate of simple single-user receivers, which can be computed efficiently for large numbers of users. A by-product of this analysis is the determination of the optimum detection threshold as a function of the number of users--a problem that finds no counterpart in the conventional Gaussian channel. We have achieved the first exact analysis of a single-user optical matched filter detector in the presence of an arbitrary number of asynchronous transmitters. The comparison of our exact analysis with popular approximations on user synchronism or on the distribution of the multiple-access interference points to severe shortcomings of those approximations, as the optimal threshold is consistently underestimated, leading to an error probability which is an order of magnitude above the one that can be obtained by a simple adjustment of the threshold as a function of the number of active transmitters.

Publications sponsored under ARO Contract DAAL-03-87-K-0062

H. V. Poor, S. Verdú, "Single-user Detectors for Multiuser Channels," *IEEE Trans. Communications,* vol. COM-36, no. 1, p. 50-60, Jan. 1988.

S. Ghez, S. Verdú, S. Schwartz, "Stability Properties of Slotted-Aloha with Multipacket Reception Capability," *IEEE Trans. Automatic Control,* vol AC-33, no. 7, p. 640-649, July 1988.

R. Lupas, S. Verdú, "Linear Multiuser Detectors for Synchronous Code Division Multiple Access Channels," *IEEE Trans. Information Theory,* vol. IT-35, no. 1, p. 123-136, Jan. 1989.

S. Verdú, "Computational Complexity of Optimum Multiuser Detection," *Algorithmica,* vol. 4, no. 3, p. 303-312, 1989.

A. Ephremides, S. Verdú, "Control and Optimization Methods in Communication Network Problems," *IEEE Trans. Automatic Control,* vol. AC-34, no. 9, p. 930-942, Sep. 1989.

S. Verdú, "Recent Progress in Multiuser Detection," *Advances in Communications and Signal Processing,* W. Porter and S. Kak, Eds., *Lecture Notes in Control and Information Sciences,* vol. 129, p. 27-38, Springer-Verlag, Heidelberg, 1989

S. Ghez, S. Verdú, S. Schwartz, "Optimal Decentralized Control in the Multipacket Channel," *IEEE Trans. Automatic Control,* vol. AC-34, no. 11, p. 1153-1163, Nov. 1989.

R. Lupas, S. Verdú, "Near-Far Resistance of Multiuser Detectors in Asynchronous Channels," *IEEE Trans. Communications,* vol. COM-38, no. 4, pp. 496-508, Apr. 1990.

D. Brady, S. Verdú, "A Semi-Classical Analysis of Optical CDMA," *IEEE Trans. Communications,* vol. COM-39, to appear.

S. Verdú, "Multiuser Detection," *Advances in Statistical Signal Processing, vol. 2. Signal Detection* H. V. Poor and J. B. Thomas, Editors, JAI Press: New York, 1990

H. V. Poor, "On Parameter Estimation in DS/SSMA Formats," *Proc. 1988 Conf. on Advances in Communications and Control Systems,* Baton Rouge, LA, Oct. 1988

R. Cheng and S. Verdú, "Capacity of RMS Bandlimited Gaussian Channels," *IEEE Trans. Information Theory,* to appear.

S. Ghez, S. Verdú and S.C. Schwartz, "Stability of Multipacket Aloha," *Proc. 21st. Conf. Information Sciences and Systems,* The Johns Hopkins University, Baltimore, MD, p. 262-267, Mar.

1987.

S. Verdú, "Fundamental Limits of Code Division Multiple Access Channels," Invited paper, *1987 IEEE Information Theory Workshop,* Bellagio, Italy, June 1987.

S. Ghez, S. Verdú and S. Schwartz, "Decentralized Control Algorithms for Multipacket Aloha," *Proc. Twenty-fifth Allerton Conf. on Communication, Control and Computing,* p. 1089-1098, Oct. 1987.

R. Lupas, S. Verdú, "Optimum Near-Far Resistance of Linear Detectors for Code-Division Multiple-Access Channels," *Abs. 1988 IFEE International Symposium on Information Theory,* Kobe, Japan, p. 14, June 1988.

S. Ghez, S. Verdú, "Ergodicity of Controlled and Uncontrolled Multipacket Aloha," *Proc. 1988 Beijing International Workshop on Information Theory,* Beijing, China, p. BII-1.1-4, July 1988.

S. Verdú, "Recent Progress in Multiuser Detection," Invited pap *1988 Int. Conf. on Advances in Communications and Control Systems,* Baton Rouge, LA, vol. 1, p. 66-77, Oct. 1988.

D. Brady and S.Verdú, "An Exact Analysis of the Optical CDMA Noncoherent Receiver," *Proc. 23rd. Conf. Information Sciences and Systems,* The Johns Hopkins University, Baltimore, MD, p. 327-332 Mar. 1989.

D. Brady and S. Verdú, "Error Probability of CDMA in the Poisson Channel," Invited paper, *1989 Workshop on Information Theory,* Cornell University, Ithaca, NY, Jun. 1989

A. Ephremides and S. Verdú, 'Control Issues in Communication Networks," *1989 IEEE Decision and Control Conference,* Tampa, Fl., Dec. 1989

R. Cheng and S. Verdú, "Capacity of RMS Bandlimited Gaussian Multiple-Access Channels," *1990 IEEE International Symposium on Information Theory* Jan. 14-19, 1990

S. Verdú, "The First Forty Years of the Shannon Theory," *Conference Record IEEE TAB Region 10 Tour,* Piscataway, NJ: IEEE Press, 1990. Also, reprinted in *IEEE Information Theory Newsletter,* Sep. 1990.

R. Cheng and S. Verdú, "On VanderMeulen's characterization of the multiacess capacity region," *1990 IEEE Information Theory Workshop,* Veldhoven, Netherlands, June 1990.

D. Brady and S. Verdú, "Semi-classical analysis of Optical CDMA receivers," *IEEE/LEOS Optical Multiple Access N tworks,* Invited paper, Monterey, CA, Jul. 1990

D. Brady and S. Verdú, "The Asymptotic Capacity of the Direct Detection Photon Channel with a Bandwidth Constraint," *Twenty-Eighth Allerton Conf. on Communication, Control and Computing,* Oct. 3-5, 1990

R. Lupas, "Near-Far Resistant Linear Multiuser Detection," Information Sciences and Systems Technical Report no. 53, (ARO) Dept. Electrical Engineering, Princeton University, Jan. 1989

S. Ghez, "Random Access Communications for the Multipacket Channel," Ph. D. Dissertation, Dept. Electrical Engineering, Princeton University, Aug. 1989

D. Brady, "Analysis of Optical, Dircet Detection Communication Systems with Point Process Observations," Ph. D. Dissertation, Dept. Electrical Engineering, Princeton University, Jul. 1990

Participating Scientific Personnel

- Principal Investigator: Sergio Verdú, Associate Professor.
- Research Associate: David Brady, Visiting Lecturer.
- Research Assistant: David Brady, Graduate Student.
- Research Assistant: Ruxandra Lupas, Graduate Student.
- Research Assistant: Smita Agarwal, Graduate Student.
- Research Assistant: Costas Diamantaras, Graduate Student.
- Consultant: H. Vincent Poor, Professor, Dept. of EE, Univ. of Illinois.

Doctoral Degrees Awarded during Contract Period

- Ruxandra Lupas, Ph.D., 1988
- Sylvie Ghez, Ph.D, 1989
- David Brady, Ph.D., 1990

T-AC, 33/7//21359

Stability Properties of Slotted Aloha with Multipacket Reception Capability

Sylvie Ghez
Sergio Verdú
Stuart C. Schwartz

Reprinted from
IEEE TRANSACTIONS ON AUTOMATIC CONTROL
Vol. 33, No. 7, July 1988

Stability Properties of Slotted Aloha with Multipacket Reception Capability

SYLVIE GHEZ, STUDENT MEMBER, IEEE, SERGIO VERDÚ, MEMBER, IEEE, AND STUART C. SCHWARTZ, SENIOR MEMBER, IEEE

Abstract—The stability of the Aloha random access algorithm in an infinite-user slotted channel with multipacket reception capability is considered. This channel is a generalization of the usual collision channel, in that it allows the correct reception of one or more packets involved in a collision. The number of successfully received packets in each slot is modeled as a random variable which depends exclusively on the number of simultaneous attempted transmissions. This general model includes as special cases channels with capture, noise, and code division multiplexing. It is shown by means of drift analysis that the channel backlog Markov chain is ergodic if the packet arrival rate is less than the expected number of packets successfully received in a collision of n as n goes to infinity. Finally, the properties of the backlog in the nonergodicity region are examined.

I. Introduction

One of the main problems in random access communications is the determination of the maximum stable throughput. In particular, an important result is that the Aloha protocol is unstable [1]–[3] in an infinite-user slotted collision channel where a transmission is successful only if no other users attempt transmissions simultaneously. Several strategies have been designed to stabilize this channel, such as collision resolution algorithms (see [4], for example) where transmissions are deferred until the current conflict is solved, and more recently, Aloha-type strategies using decentralized control, where the retransmission probability is updated according to previous channel outcomes. It has been shown [5]–[7] that the maximum stable throughput achievable by such Aloha-type strategies with decentralized control is e^{-1}.

However, the collision channel model does not hold in many important practical multiuser communication systems [8]–[21] because simultaneous transmission of several packets does not necessarily result in the destruction of all the transmitted information. For instance, the capture phenomenon is common in local area radio networks [12]–[15]; if the power of one of the received packets is sufficiently large compared to the power of the other packets involved in a collision, then the strongest packet can be correctly decoded, while the other packets are lost. Other examples are multiple-access channels where several users transmit simultaneously in the same frequency band, and a multiuser detector demodulates the information transmitted by all active users (e.g., [8]–[11]). Although those systems do not necessarily require a random access protocol, it is sometimes useful to exercise some flow control through such a protocol so as to limit the maximum number of simultaneous transmitters, in order to bound the multiuser receiver complexity and guarantee lower bit-error rates.

Manuscript received July 23, 1987; revised January 8, 1988. Paper recommended by Past Associate Editor A. Ephremides. This work was supported in part by the Office of Naval Research under Contract N00014-87-k-0054 and by the Army Research Office under Contract DAAL03-87-k-0062.

The authors are with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544.

IEEE Log Number 8821359.

Previous studies of some of the aforementioned systems [9], [12]–[18] where some of the packets involved in a collision may be correctly received have shown that the performances are noticeably improved with respect to slotted Aloha. However, even in those special cases, no precise stability result is available, either because finite population networks with no buffer space were considered, or because the Poisson approximation of channel traffic was used for infinite population networks. In [19] (see also [20]), upper and lower bounds are derived for the capacity of a multiple access channel where all packets are correctly received if the collision size does not exceed a fixed threshold and otherwise all packets are destroyed.

In this paper, we consider a generalization of the collision channel, where the receiver can demodulate several packets simultaneously. It is assumed that the number of correctly demodulated packets is a random variable, which, given the number of packets simultaneously transmitted, is independent of the backlog and of the number of previous retransmission attempts. This random variable can take any integer value between zero and the collision size. Thus, the channel is described by a matrix of conditional probabilities (ϵ_{nk}) where ϵ_{nk} is the probability that k packets are correctly demodulated given that there were n simultaneous transmissions. We analyze the usual Aloha algorithm with the multipacket reception capability just described. Users are synchronized so that transmissions take place within one slot, and at the end of each slot, stations that did transmit a packet learn whether or not their transmission was successful. Unsuccessful or backlogged packets are retransmitted in each subsequent slot with probability p, $0 < p \leq 1$. It turns out that multipacket reception capability can stabilize Aloha. Our main result states that the maximum stable throughput is equal to the limit of the average number of packets correctly received in collisions of size n when n goes to infinity. To show this, we model the channel backlog as a Markov chain, and then study its properties by using some simple drift analysis techniques.

The last part of this paper is a study of the properties of the backlog in the nonergodicity region. Unlike the backlog Markov chain for slotted Aloha which is always transient [1], the backlog for our model does in general have a null recurrence region of positive length, which depends on the matrix (ϵ_{nk}) and on the retransmission probability p. However, transience in the nonergodicity region can be ensured for a large class of systems, and in particular for channels where the number of successful simultaneous transmissions is bounded.

II. Multipacket Reception Model

Let A_k be the number of new packets arriving during time slot k. Assume that $(A_k)_{k\geq 0}$ are i.i.d. random variables with probability distribution:

$$P[A_k = n] = \lambda_n \qquad (n \geq 0)$$

such that the mean arrival rate $\lambda = \sum_{n=1}^{\infty} n\lambda_n$ is finite. New packets are transmitted with probability one at the beginning of the first slot following their arrival.

Given that n packets are being transmitted in one slot, we define for $n \geq 1, 0 \leq k \leq n$

$$\epsilon_{nk} = P[k \text{ packets are correctly received} \mid n \text{ are transmitted}].$$

The multipacket reception properties of the channel are summarized by the stochastic matrix

$$E = \begin{bmatrix} \epsilon_{10} & \epsilon_{11} & & \\ \epsilon_{20} & \epsilon_{21} & \epsilon_{22} & 0 \\ \cdot & \cdot & & \\ \epsilon_{n0} & \epsilon_{n1} & & \epsilon_{nn} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & & \cdot \end{bmatrix}$$

which we refer to as the *reception matrix* of the channel. For instance, the reception matrix for the usual collision channel is

$$\begin{bmatrix} 0 & 1 & \\ 1 & 0 & 0 \\ 1 & 0 & \\ 1 & \cdot & \\ \cdot & \cdot & \\ \cdot & \cdot & \end{bmatrix}$$

while for a system with capture it has the form

$$\begin{bmatrix} 0 & 1 & \\ 1-x_2 & x_2 & 0 \\ \cdot & \cdot & \\ 1-x_n & x_n & \\ \cdot & \cdot & \\ \cdot & \cdot & \end{bmatrix}$$

where x_n is the probability of capture given that the collision size is n. The model studied in [19], [20] can be described by a reception matrix of the form

$$\begin{bmatrix} 0 & 1 & & \\ 0 & 0 & 1 & 0 \\ \cdot & \cdot & & \cdot \\ 0 & 0 & & 1 \\ 1 & 0 & & \\ 1 & 0 & & 0 \\ \cdot & \cdot & & \end{bmatrix}.$$

Note that by letting $\epsilon_{10} \neq 0$ our model allows not only collisions but also background noise to be a source of errors.

Denote by X_n the number of backlogged packets in the system at the beginning of slot n. The discrete-time process $(X_n)_{n \geq 0}$ is easily seen to be a homogeneous Markov chain. We define the system to be stable if $(X_n)_{n \geq 0}$ is ergodic and unstable otherwise. The average number of packets correctly received in collisions of size n is denoted by $C_n = \sum_{k=1}^{n} k\epsilon_{nk}$. We can now state the main result.

Theorem 1: If C_n has a limit $C = \lim_{n\to\infty} C_n$, then[1] the system is stable for all arrival distributions such that $\lambda < C$ and is unstable for $\lambda > C$. This also holds if C is infinite. if $\lim_{n\to\infty} C_n = +\infty$, then the system is always stable.

The proof is given in Section III. In the remainder of this section, we use Theorem 1 to analyze several simple random access channels that fall within the scope of the multipacket reception channel.

1) Mobile Users with Pairwise Transmissions. Consider an infinite number of transmitters $T_1, T_2, \cdots$, and an infinite number of receivers $R_1, R_2, \cdots$, whose positions in the plane are i i d random variables. Suppose that transmissions are pairwise

[1]This result holds under the assumption that the Markov chain of the number of backlogged packets is irreducible and aperiodic (for details and sufficient conditions, see Section III).

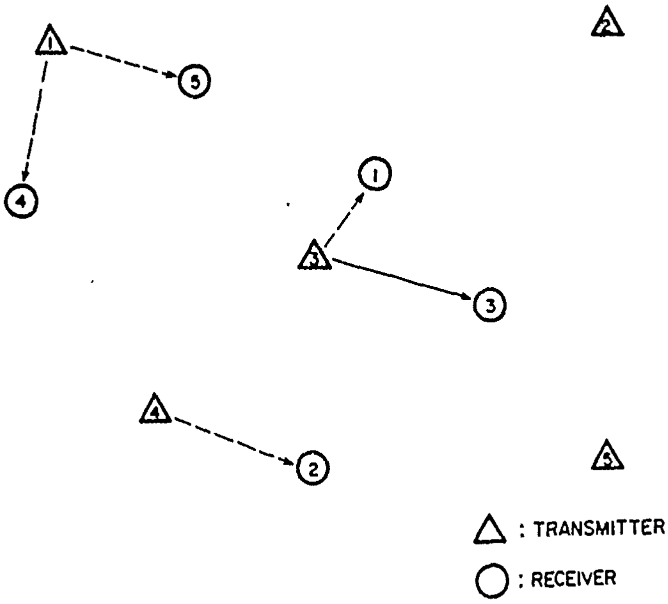


Fig. 1. Pairwise transmissions with only one success (3-3).

in the sense that transmitter T_n sends packets only to receiver R_n, and R_n is only interested in the packets sent by T_n (see Fig. 1). Assume also that each receiver can only detect correctly the packet sent by the closest transmitter (in particular, this is the case if there is perfect capture, see Example 3 below). The successes of transmissions occurring at the same time are independent, so that for $n \geq 2$

$$\epsilon_{nk} = \binom{n}{k} p(n)^k (1-p(n))^{n-k}$$

where $p(n)$ is the probability that any given transmitter is successful in a collision of size n, which is equal to $1/n$ if we assume that all locations are memoryless, i.e., independent from slot to slot. It follows that

$$C_n = np(n) = 1$$

and the maximum throughput is 1. More generally, if because of channel noise, the message of the closest transmitter is received correctly with probability α (in other words $\epsilon_{11} = \alpha$), then the throughput is equal to α. The assumption that the locations of the stations are memoryless is equivalent to assuming that they move infinitely fast If this simplifying assumption is dropped, then the number of successes depends not only on the current number of retransmissions, but also on the previous history of retransmissions, and thus the problem is no longer encompassed by our multipacket reception model. In Fig. 2, the result of a simulation shows that for moderate speeds, the actual throughput is well approximated by the foregoing analysis.

2) Frequency Hopping Random Access Channel: Consider a finite population of N users transmitting by frequency hopping, as in [11], [22]. For each packet he wants to transmit, a user selects with equal probability one frequency in a fixed set of q frequencies. A packet is correctly received iff no other packet is transmitted on the same frequency during the same slot. We compute $(\epsilon_{Nk})_{1 \leq k \leq N}$, and $C = \lim_{N\to\infty} C_N$. If the users have infinite buffer space, then C can be taken as a good approximation for large N of the maximum stable throughput of the system, which is unknown. If the users have no buffer space, as is often assumed, the backlog Markov chain is always ergodic, but even then, one should expect reasonable delays in large population problems only for arrival rates below C. The computation of the reception matrix of this channel is a simple combinatorial problem of random assignment of objects to cells (e.g., see [23, App. A]).

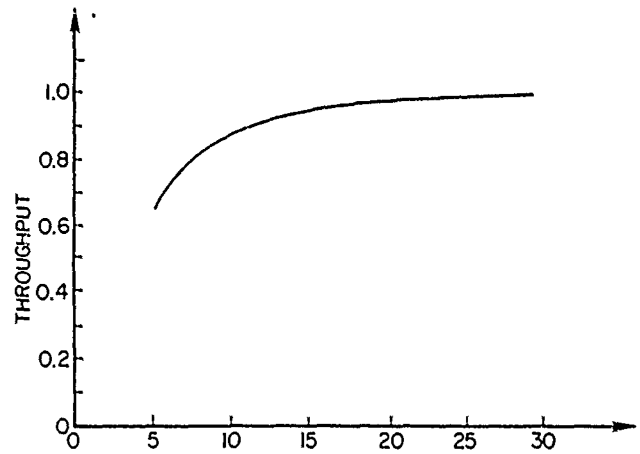


Fig. 2. Throughput as a function of velocity for mobile users with pairwise transmissions. Stations moving in a square region; velocity units: percentage of square side traveled in one slot. Retransmission probability set to 0.1.

Denote by $T_1, T_2, \cdots, T_N$ the users, all involved in the collision, and also denote by S the set of users whose packets are correctly received. Two cases need to be considered.

a) $2 \le N \le q$: We have, for $1 \le j \le N$

$$\epsilon_{Nj} = \binom{N}{j} P[S=\{T_1, T_2, \cdots, T_j\}] \tag{1}$$

and the following decomposition:

$$P[\{T_1, T_2, \cdots, T_j\} \subseteq S] = P[S=\{T_1, T_2, \cdots, T_j\}] + P\left[\bigcup_{k=j+1}^{N} \{\{T_1, \cdots, T_j, T_k\} \subseteq S\}\right]$$

easily yields the desired expression

$$P[S=\{T_1, T_2, \cdots, T_j\}] = \sum_{k=0}^{N-j} (-1)^k \cdot \binom{N-j}{k} P[\{T_1, T_2, \cdots, T_{k+j}\} \subseteq S] \tag{2}$$

where only one term is left to compute

$$P[\{T_1, T_2, \cdots, T_{k+j}\} \subseteq S] = \frac{q(q-1) \cdots (q-j-k+1)(q-j-k)^{N-j-k}}{q^N} \tag{3}$$

for $1 \le j \le N$, $0 \le k \le N-j$. Putting (1), (2), and (3) together gives the result

$$\epsilon_{Nj} = \binom{N}{j} \sum_{k=0}^{N-j} (-1)^k \binom{N-j}{k} \cdot \frac{q(q-1) \cdots (q-j-k+1)(q-j-k)^{N-j-k}}{q^N} \tag{4}$$

for $1 \le j \le N$. Notice in particular that $\epsilon_{N,N-1} = 0$. Let us now compute the average number of packets correctly received in collisions of size N, $C_N = \sum_{j=1}^{N} j\epsilon_{Nj}$. By using (4) and summing at $j + k$ constant, we get

$$q^N C_N = \sum_{i=1}^{N} q(q-1) \cdots (q-i+1)(q-i)^{N-i} \cdot \sum_{n=0}^{i-1} (-1)^n \frac{N!}{n!(i-n-1)!(N-i)!}$$

which can be simplified as

$$q^N C_N = \sum_{i=1}^{N} \frac{N!}{(i-1)!(N-i)!} \cdot q(q-1) \cdots (q-i+1)(q-i)^{N-i}(1-1)^{i-1}$$

to get the final result

$$C_N = N\left(1 - \frac{1}{q}\right)^{N-1}.$$

b) $N > q$: In this case, there can be at best $q-1$ successes in a collision of size N. The same method applies to get the following probabilities:

$$\epsilon_{Nj} = \binom{N}{j} \sum_{k=0}^{q-j-1} \binom{N-j}{k} (-1)^k \cdot \frac{q(q-1) \cdots (q-j-k+1)(q-j-k)^{N-j-k}}{q^N} \quad (1 \le j \le q-1)$$

$$\epsilon_{Nj} = 0 \quad (q \le j \le N)$$

resulting in the same expected number of successes as before

$$C_N = N\left(1 - \frac{1}{q}\right)^{N-1}.$$

Now we let the population size N go to infinity and we apply our result. If we let N grow to infinity while keeping q constant, we have $\lim_{N\to\infty} C_N = 0$, so the system is always unstable. On the other hand, if we let N go to infinity while keeping q equal to a fixed percentage of the population size, i.e., N/q constant, then $\lim_{N\to\infty} C_N = +\infty$, and the system is always stable. It is easily shown that to get a finite maximum stable throughput, q has to grow as $N/\ln N$.

3) Mobile Radio Network with Capture: Consider an infinite number of users independently and uniformly distributed in a circle of radius R, whose positions are independent from slot to slot. Users transmit packets to a common receiver located at the center of the network. Denote by P_1 and P_2 the received powers of the strongest and the next to strongest packets involved in a collision. Assume, as in [12]–[14], that the strongest packet is correctly received iff $P_1/P_2 > K$ (K being a system dependent constant), and that all the other packets involved in the collision are not received successfully. Assume, moreover, that the received power of a packet only depends on the distance r between the sender and the receiver

$$P = \frac{\text{constant}}{r^\alpha} \quad (\alpha \ge 2).$$

Then there will be capture iff

$$r_2 > \beta r_1$$

where $\beta = K^{1/\alpha}$ is the capture parameter, and r_1, r_2 are the distances of the closest and the next to closest senders from the receiver.

Denote by D the distance between a given user and the receiver. It is easily shown that the pdf of D is given by

$$p_D(r)=2\frac{r}{R^2} \qquad (0\le r\le R).$$

Given N users, denote by U_N the closest from the receiver, and by D_N its distance from the receiver. Computing the cdf of D_N and taking its derivative, we obtain

$$p_{D_N}(r)=2N\frac{r}{R^2}\left[1-\left(\frac{r}{R}\right)^2\right]^{N-1} \qquad (0\le r\le R). \tag{5}$$

Given $D_N = r$, the other $N-1$ users are uniformly distributed in the annular region (r, R). So if N users collide and $D_N = r$, U_N will be correctly received iff all the other users are in the annular region $(\beta r, R)$, which is empty if $\beta r > R$. Therefore, if we denote by

$$P_N(r)=P[\text{capture}|N \text{ collide}, D_N=r] \qquad (N\ge 2)$$

we have

$$P_N(r)=\begin{cases}\left[\dfrac{R^2-\beta^2r^2}{R^2-r^2}\right]^{N-1} & \text{if } r\le\dfrac{R}{\beta}\\ 0 & \text{if } r\ge\dfrac{R}{\beta}\end{cases}. \tag{6}$$

Thus, the probability of capture in a collision of N ($N \ge 2$) is

$$\epsilon_{N1}=\int_0^{R/\beta} P_N(r)p_{D_N}(r)\,dr.$$

Using (5) and (6), and with the change of variable $x = \beta/R$, this is easily computed

$$\epsilon_{N1}=\int_0^{1/\beta} 2Nx(1-\beta^2x^2)^{N-1}\,dx=\frac{1}{\beta^2}.$$

It follows that $C = 1/\beta^2$ is the maximum stable throughput. Notice, in particular, that for $\beta = 1$ (perfect capture), we have $C = 1$ and for $\beta \to \infty$ (no capture), we have $C \to 0$.

Under certain conditions, the performances of Aloha in the multipacket channel can be improved by varying the retransmission probability as a function of the channel history, and a maximum stable throughput of $\sup_{x\ge 0} e^{-x}\sum_{n=1}^{\infty} C_n/n!x^n$ can be reached (see [31]).

III. Ergodicity Region

The number of backlogged packets in the system at time n, $(X_n)_{n\ge 0}$, is a homogeneous Markov chain whose one-step transition probability matrix can be computed as a function of p, $(\lambda_k)_{k\ge 0}$, and E. Denoting by $B_i(j)$ the probability of having j retransmissions out of i backlogged packets

$$B_i(j)=\binom{i}{j}p^j(1-p)^{i-j} \tag{7}$$

we get

$$P_{00}=\lambda_0+\sum_{n=1}^{\infty}\lambda_n\epsilon_{nn}$$

$$P_{0k}=\sum_{n=0}^{\infty}\lambda_{k+n}\epsilon_{k+n,n} \qquad (k\ge 1)$$

and for $i \ge 1$

$$P_{i,i-k}=\sum_{n=0}^{\infty}\lambda_n\sum_{j=k}^{i}B_i(j)\epsilon_{n+j,n+k} \qquad (1\le k\le i)$$

$$P_{ii}=\lambda_0\left[B_i(0)+\sum_{j=1}^{i}B_i(j)\epsilon_{j0}\right]+\sum_{n=1}^{\infty}\lambda_n\sum_{j=0}^{i}B_i(j)\epsilon_{n+j,n}$$

$$P_{i,i+k}=\sum_{n=0}^{\infty}\lambda_{k+n}\sum_{j=0}^{i}B_i(j)\epsilon_{j+k+n,n} \qquad (k\ge 1). \tag{8}$$

Sufficient conditions for $(X_n)_{n\ge 0}$ to be irreducible and aperiodic are as follows:

- if $0 < p < 1$:

$$\lambda_0\ne 0 \tag{9a}$$

$$\lambda_0+\sum_{n=1}^{\infty}\lambda_n\epsilon_{nn}<1 \tag{9b}$$

$$\epsilon_{10}\ne 1 \tag{9c}$$

- if $p = 1$:

$$\lambda_0\ne 0 \tag{9a}$$

$$\lambda_0+\sum_{n=1}^{\infty}\lambda_n\epsilon_{nn}<1 \tag{9b}$$

$$\text{for all } i\ge 1,\ \epsilon_{i0}\ne 1. \tag{9d}$$

These are only sufficient conditions, but they hold for almost all nontrivial systems. For example, if (9b) does not hold, then zero is an absorbing state, since the left-hand side of (9b) is equal to P_{00}. Also, (9c) simply means that the successful reception of a single packet in the absence of other active users is possible. Assume, for instance, that $0 < p < 1$ and that the arrivals are Poisson distributed. Then we only have to assume (9c), and (9b) is true unless there is perfect reception, that is $\epsilon_{nn} = 1$ for all $n \ge 1$, in which case the system would of course always be stable. The case $p = 1$ gives rise to a number of pathological situations, hence, the much stronger condition (9d). It generally turns out that either (9d) is not necessary or the stability region of the system is obvious. For instance, it is clear from the transition probabilities that slotted Aloha with $p = 1$ is always unstable. In any case, it is assumed in what follows that $(X_n)_{n\ge 0}$ is irreducible and aperiodic.

Proof of Theorem 1: The proof is based on drift analysis. Recall that in general, the drift at state i ($i \ge 0$) is defined by

$$d_i=E[X_{t+1}-X_t|X_t=i].$$

If we denote by Σ_t the number of successful transmissions in slot t, we have

$$X_{t+1}-X_t=A_t-\Sigma_t$$

and therefore

$$d_i=\lambda-E[\Sigma_t|X_t=i]. \tag{10}$$

Now if R_t is the number of retransmissions in slot t, we get

$$P[\Sigma_t=k|X_t=i,\ A_t=n,\ R_t=j]=\epsilon_{n+j,k}$$

for $0 \le j \le i$, $0 \le k \le n + j$ and with the convention that $\epsilon_{00} =$

$C_0 = 0$. Thus,

$$E[\Sigma_t | X_t = i,\ A_t = n,\ R_t = j] = C_{n+j}$$

and

$$E[\Sigma_t | X_t = i] = \sum_{n=0}^{\infty} \lambda_n \sum_{j=0}^{i} B_i(j) C_{n+j}. \tag{11}$$

The value of the drifts for our model follows from (10) and (11)

$$d_i = \lambda - \sum_{n=0}^{\infty} \lambda_n \sum_{j=0}^{i} B_i(j) C_{n+j}. \tag{12}$$

The idea of the proof is to compute $\lim_{i\to\infty} d_i$ which will turn out to be a very simple expression, and then apply the results of [3] and [24] to determine the ergodicity region of $(X_n)_{n\geq 0}$. Let us first recall the two results that will be used in the sequel.

Lemma A (Pakes [24]): Let $(X_n)_{n\geq 0}$ be an irreducible and aperiodic Markov chain having as state space the nonnegative integers, denote by (P_{ij}) its transition probability matrix, and by d_i its drift at state i. Then if for all i $|d_i| < \infty$, and if $\limsup_{i\to\infty} d_i < 0$, $(X_n)_{n\geq 0}$ is ergodic.

Lemma B (Kaplan [3]): Under the assumptions of Lemma A, if for some integer $N \geq 0$ and some constants $B \geq 0$, $c \in [0, 1]$ the following two conditions hold, then $(X_n)_{n\geq 0}$ is not ergodic:

i) for all $i \geq N$, $d_i > 0$

ii) for all $i \geq N$, all $\theta \in [c, 1]$, $\theta^i - \Sigma_j P_{ij}\theta^j \geq -B(1-\theta)$.

From (12), it can be seen that $|d_i|$ is finite since

$$|d_i| \leq \lambda + \sum_{n=0}^{\infty} \lambda_n \sum_{j=0}^{i} B_i(j) C_{n+j} \leq 2\lambda + ip.$$

Next, the drift limit is given by the following lemma.

Lemma 1: If C_n has a limit C, finite or not, then $\lim_{i\to\infty} \Sigma_{n=0}^{\infty} \lambda_n \Sigma_{j=0}^{i} B_i(j) C_{n+j} = C$.

Proof of Lemma 1: We consider two separate cases depending on whether C is finite.

1) $C = +\infty$.

Fix $A > 0$ and pick $r \geq 0$ such that $\lambda_r \neq 0$. There exists an integer M such that for all $n \geq M$, $C_n > A$. Fix such an M. Then we have for $i \geq M$

$$\sum_{n=0}^{\infty} \lambda_n \sum_{j=0}^{i} B_i(j) C_{n+j} > \lambda_r \sum_{j=0}^{i} B_i(j) C_{j+r} > \lambda_r A \sum_{j=M}^{i} B_i(j)$$

which terminates the proof, since for any fixed $M \geq 0$

$$\lim_{i\to\infty} \sum_{j=M}^{i} B_i(j) = 1. \tag{13}$$

2) $C < +\infty$.

We have for $i > M$

$$\left| \sum_{n=0}^{\infty} \lambda_n \sum_{j=0}^{i} B_i(j) C_{n+j} - C \right| \leq \sum_{j=0}^{M} B_i(j) \sum_{n=0}^{\infty} \lambda_n |C_{n+j} - C| + \sum_{j=M+1}^{i} B_i(j) \sum_{n=0}^{\infty} \lambda_n |C_{n+j} - C|. \tag{14}$$

Fix $\epsilon > 0$. There exists M such that for all $n > M$, $|C_n - C| < \epsilon/2$. Fix such an M. Then

$$\sum_{j=M+1}^{i} B_i(j) \sum_{n=0}^{\infty} \lambda_n |C_{n+j} - C| < \frac{\epsilon}{2}.$$

Also, if L is an upper bound for C_n

$$\sum_{j=0}^{M} B_i(j) \sum_{n=0}^{\infty} \lambda_n |C_{n+j} - C| \leq 2L \sum_{j=0}^{M} B_i(j) < \frac{\epsilon}{2}$$

for i big enough because (13) holds, which takes care of the first term in (14) and ends the proof of Lemma 1.

Putting together (12) and Lemmas A and 1, we get that 1) if $\lim_{n\to\infty} C_n = +\infty$, then $\lim_{i\to\infty} d_i = -\infty$, and $(X_n)_{n\geq 0}$ is ergodic; and 2) if $\lim_{n\to\infty} C_n = C < +\infty$, then $\lim_{i\to\infty} d_i = \lambda - C$, and $(X_n)_{n\geq 0}$ is ergodic for $\lambda < C$. If $\lambda > C$, we can apply Lemma B and conclude that $(X_n)_{n\geq 0}$ is not ergodic provided that Kaplan's condition ii) holds. This is the purpose of Lemma 2, which is the last step in the proof of Theorem 1.

Lemma 2: If for all $n \geq 1$, $C_n < L$ for some $L \in (0, \infty)$, then Kaplan's condition holds: there exists a constant B, an integer N, and a real $c \in [0, 1]$ such that

$$\theta^i - \sum_j P_{ij}\theta^j \geq -B(1-\theta) \qquad \text{all } i \geq N,\ \theta \in [c, 1].$$

Proof of Lemma 2: According to [25], it is enough to show that the downward part of the drift, defined as

$$D(i) = -\sum_{k=1}^{i} k P_{i,i-k}$$

is bounded below. From the transition probabilities (8), we get

$$D(i) = -\sum_{k=1}^{i} k \sum_{n=0}^{\infty} \lambda_n \sum_{j=k}^{i} B_i(j) \epsilon_{n+j,n+k}$$

which can also be put in the form

$$D(i) = -\sum_{j=1}^{i} B_i(j) \sum_{n=0}^{\infty} \lambda_n \sum_{k=1}^{j} k \epsilon_{n+j,n+k}$$

from which it follows that

$$D(i) \geq -\sum_{j=1}^{i} B_i(j) \sum_{n=0}^{\infty} \lambda_n C_{n+j} \geq -L.$$

□

Notice that in the proof of Theorem 1 (and this also holds for Theorem 2 below), the exact expression (7) for $B_i(j)$ is never used. The only requirements are that $(B_i(j))_{0\leq j\leq i}$ is a probability distribution, and that (13) holds. Therefore, our results are valid for a larger class of retransmission policies than was first assumed. For example, there could be K priority groups, each with a different retransmission probability.

Although Theorem 1 is quite general, in many practical cases, the reception matrix has a very simple structure and the stability region can be obtained with virtually no computations. This happens for instance in radio networks with capture where all is needed is the limit of the second column of the matrix, or also in the simple case where above a certain collision size N, the transmission is too garbled for the receiver to be able to decode anything correctly, so that $C_n = 0$ for $n > N$.

This last example is a particular case of a noteworthy feature of Theorem 1, namely that the stability region does not depend on any finite number of rows of the reception matrix. In fact, any number of modifications of the matrix that leaves $\lim_{n\to\infty} C_n$ unchanged does not affect the stability region. Although this may be surprising at first sight, it can be intuitively explained by the fundamental instability of the collision channel: unless the receiver is perfect (all ϵ_{nn} equal to 1), the backlog will eventually exceed any prefixed value with probability one, thus it is the limit of C_n that determines the stability region.

The stability region is also unchanged if the first transmission of packets is delayed. If new packets are backlogged, that is, transmitted for the first time with probability p in each slot following their arrival (this transmission rule appears in the literature as controlled-access or delayed first transmission), the drifts become $d_i = \lambda - \sum_{j=1}^{i} B_i(j)C_j$ for $i \geq 1$, and from Lemmas 1 and 2 the ergodicity region remains the same.

If C_n does not have a limit, Theorem 1 does not give the stable throughput of the system. Even though in almost all practical cases, and indeed in all the examples of Section II, C_n does have a limit, it is conceptually interesting to examine the case when $\liminf_{n\to\infty} C_n \neq \limsup_{n\to\infty} C_n$. It is worth pointing out that adding constraints as strong as the following on the reception matrix still does not imply that C_n has a limit:

i) $(\epsilon_{n0})_{n\geq 1}$ is nondecreasing

ii) $(\epsilon_{nk})_{n\geq k}$ is nonincreasing for all $k \geq 1$

iii) $\epsilon_{nk} \geq \epsilon_{n,k+1}$ for $n \geq 2$, $1 \leq k \leq n-1$

although the counterexamples we have been able to build are somewhat contrived. Notice that conditions i) and ii) above imply that each column has a limit $\alpha_k = \lim_{n\to\infty} \epsilon_{nk} (k \geq 0)$, which is very likely to happen in practice. In any case, Theorem 2 below still gives some information on the stability region, although the exact result requires in general the complete knowledge of the sequence $(C_n)_{n\geq 1}$. In fact, given any nonnegative numbers $\alpha < \gamma < \beta$, one can construct a reception matrix with nth row average C_n such that:

i) $\liminf_{n\to\infty} C_n = \alpha$

ii) $\limsup_{n\to\infty} C_n = \beta$

and such that the maximum stable throughput is γ.

Theorem 2: The system is stable for $\lambda < \liminf_{n\to\infty} C_n$ and unstable for $\lambda > \limsup_{n\to\infty} C_n$.

Proof:

a) If $\lambda < \liminf_{n\to\infty} C_n$, then $(X_n)_{n\geq 0}$ is ergodic.

If $\liminf_{n\to\infty} C_n = +\infty$, then $\lim_{n\to\infty} C_n = +\infty$, and the result has already been proved, so assume that $\liminf_{n\to\infty} C_n$ is finite From Lemma A, it is enough to prove that for all $\epsilon > 0$, there exists N such that

$$d_i < \lambda - \liminf_{n\to\infty} C_n + \epsilon \qquad \text{all } i \geq N.$$

Recall from (12) that we have

$$d_i = \lambda - \sum_{n=0}^{\infty} \lambda_n \sum_{j=0}^{i} B_i(j) C_{n+j}. \tag{15}$$

So it is only needed to prove that for all $\epsilon > 0$ there exists N such that

$$\sum_{n=0}^{\infty} \lambda_n \sum_{j=0}^{i} B_i(j) C_{n+j} > \liminf_{n\to\infty} C_n - \epsilon \qquad \text{all } i \geq N.$$

Now by definition there exists M such that for all $k \geq M$:

$$C_k > \liminf_{n\to\infty} C_n - \epsilon$$

and therefore for all $i > M$:

$$\sum_{n=0}^{\infty} \lambda_n \sum_{j=0}^{i} B_i(j) C_{n+j} > (\liminf_{n\to\infty} C_n - \epsilon) \sum_{j=M}^{i} B_i(j)$$

which completes the proof since (13) holds.

b) If $\lambda > \limsup_{n\to\infty} C_n$, then $(X_n)_{n\geq 0}$ is not ergodic.

Since λ is finite, in this case $\limsup_{n\to\infty} C_n$ is necessarily finite. Therefore, $(C_n)_{n\geq 1}$ is bounded and from Lemma 2, Kaplan's condition holds. Thus, it is enough to show that for all $\epsilon > 0$, there exists N such that

$$d_i > \lambda - \limsup_{n\to\infty} C_n - \epsilon \qquad \text{all } i \geq N.$$

From (15), we only need to show

$$\sum_{n=0}^{\infty} \lambda_n \sum_{j=0}^{i} B_i(j) C_{n+j} < \limsup_{n\to\infty} C_n + \epsilon \qquad \text{all } i \geq N.$$

Since there exists M such that for all $k \geq M$

$$C_k < \limsup_{n\to\infty} C_n + \epsilon$$

then if L is an upper bound for C_n, we have for $i \geq M$

$$\sum_{n=0}^{\infty} \lambda_n \sum_{j=0}^{i} B_i(j) C_{n+j} < L \sum_{j=0}^{M-1} B_i(j) + \limsup_{n\to\infty} C_n + \epsilon$$

from which the result follows, using (13). □

IV. Behavior of the Backlog Markov Chain in the Nonergodicity Region

In this section, we further investigate the properties of $(X_n)_{n\geq 0}$ in the case $\lambda > C$, assuming of course that $(C_n)_{n\geq 1}$ has a finite limit. It has been proved in [1] that the backlog Markov chain for the usual slotted Aloha algorithm is transient, but this result cannot be generalized to our model when $\lambda > C$. We give below an example showing that $(X_n)_{n\geq 0}$ can be null recurrent when the mean arrival rate λ belongs to an interval of positive length. The boundary between the null recurrence and the transience regions generally depends in a rather complicated manner on both the reception matrix and the retransmission probability p. We give a sufficient condition for $(X_n)_{n\geq 0}$ to be transient when $\lambda > C$, as well as bounds on the recurrence region.

Consider the reception matrix defined by

$$\epsilon_{nk} = \frac{1}{n^2} \qquad (1 \leq k \leq n)$$

$$\epsilon_{n0} = 1 - \frac{1}{n}$$

for $n \geq 1$. Then $C_n = \sum_{k=1}^{n} k/n^2 = (n+1)/2n$, and $C = 1/2$. Using Lemmas C and D below, we show in [26] that X_n is recurrent for $\lambda < R(p)$ and transient for $\lambda > R(p)$, where $R(p)$ is a function of the retransmission probability p and is given by

$$R(p) = \frac{1}{p} + \frac{(1-p)}{p^2} \ln(1-p) \qquad (0<p<1)$$

$$R(1) = 1.$$

It is easily seen that $R(p)$ is an increasing function of p for $p \in]0, 1[$ with extrema $\lim_{p\to 0} R(p) = 1/2$ and $\lim_{p\to 1} R(p) = 1$. Fig. 3 summarizes the behavior of X_n for this example.

It is somehow surprising to see that in this case, as well as in all

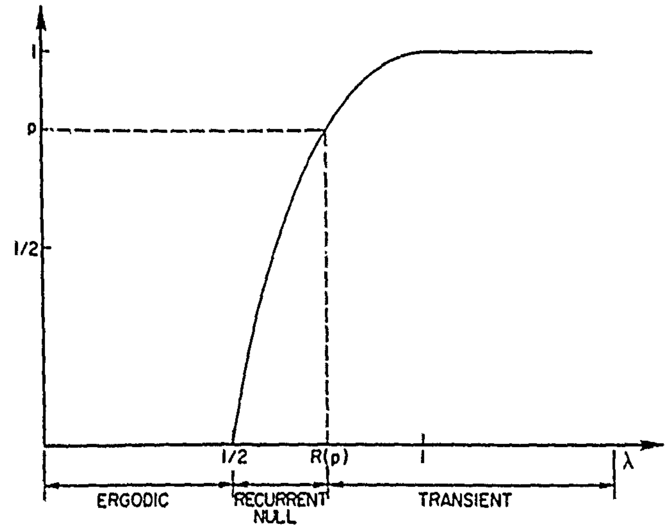


Fig. 3. Transience and ergodicity regions as a function of the retransmission probability when $\epsilon_{nk} = 1/n^2$.

the other examples we have computed, the recurrence region becomes larger as p increases. Intuitively, the recurrence of X_n when $\lambda > C$ seems to be due to the fact that transitions from any state i to 0 (or to some fixed integer k_0) are possible and that the probability of such an event, P_{i0} (or P_{ik_0}), goes to zero slowly with i. It can be checked that these probabilities are increasing functions of p when i is large enough.

Transience is ensured for $\lambda > C$ if the supremum of the elements of the kth-column goes to zero faster than k^2. This condition holds for all the examples in Section II, as well as for many real life cases, due to the practical limitations on the receiver capabilities. In particular, it is always verified if the reception matrix has only a finite number of nonzero columns (or equivalently, if the backlog Markov chain has uniformly bounded downwards transitions, as defined in [3]) which happens, for instance, if there is capture. Note that the proof of Theorem 3 below is of course valid for the conventional collision channel, and in this case becomes somewhat simpler than the proof in [1].

Theorem 3: If $\lim_{k\to\infty} k^2 \sup_{n\geq k} \epsilon_{nk} = 0$, then $(X_n)_{n\geq 0}$ is transient for $\lambda > C$.

Because of the complexity and lack of structure of the one-step transition probabilities (8), few results on the recurrence and transience of Markov chains can be applied to our model. Before proving Theorem 3, let us introduce the following two criteria from [27].

Lemma C: Let $(X_n)_{n\geq 0}$ be an irreducible and aperiodic Markov chain, having as state space the set of nonnegative integers, and with one-step transition probability matrix $P = (P_{ij})$. $(X_n)_{n\geq 0}$ is recurrent if and only if there exists a sequence $(y_n)_{n\geq 0}$ such that

1) $\lim_{n\to\infty} y_n = +\infty$

2) for some integer $N > 0$ $\sum_{j=0}^{\infty} y_j P_{ij} \leq y_i$ all $i \geq N$.

We will only use the sufficiency part, which has also been proved in [24].

Lemma D: With the same assumptions as in Lemma C, $(X_n)_{n\geq 0}$ is transient if and only if there exists a sequence $(y_n)_{n\geq 0}$ such that

1) $(y_n)_{n\geq 0}$ is bounded

2) for some integer $N > 0$ $\sum_{j=0}^{\infty} y_j P_{ij} \leq y_i$ all $i \geq N$

3) for some $k \geq N$ $y_k < y_0, \cdots, y_{N-1}$.

Sufficiency under the additional constraints $y_i > 0$ and $\lim_{i\to\infty} y_i = 0$ has also been proved in [28]. Also, the sufficiency parts of both lemmas are an immediate consequence of [29, Theorems 5 and 6] together with the results in [30].

Proof of Theorem 3: We use Lemma D with $y_n = 1/(n+1)^\theta$, $\theta \in]0, 1[$. We have

$$\sum_j P_{ij} y_j \leq y_i \Leftrightarrow \sum_{k=1}^{i} (y_{i-k} - y_i) P_{i,i-k} + \sum_{k=1}^{\infty} (y_{i+k} - y_i) P_{i,i+k} \leq 0 \quad (16)$$

and

$$(i+1)^{1+\theta} \sum_{k=1}^{i} (y_{i-k} - y_i) P_{i,i-k} + (i+1)^{1+\theta} \cdot \sum_{k=1}^{\infty} (y_{i+k} - y_i) P_{i,i+k} = D'(i) + U'(i) \quad (17)$$

where we have defined

$$D'(i) = (i+1)^{1+\theta} \sum_{k=1}^{i} \left[\frac{1}{(i+1-k)^\theta} - \frac{1}{(i+1)^\theta} \right] \cdot \sum_{n=0}^{\infty} \lambda_n \sum_{j=k}^{i} B_i(j) \epsilon_{n+j,n+k}$$

$$U'(i) = (i+1)^{1+\theta} \sum_{k=1}^{\infty} \left[\frac{1}{(i+1+k)^\theta} - \frac{1}{(i+1)^\theta} \right] \cdot \sum_{n=0}^{\infty} \lambda_{k+n} \sum_{j=0}^{i} B_i(j) \epsilon_{n+k+j,n}. \quad (18)$$

The drift of X_n at state i can be computed from the transition probabilities (8)

$$d_i = -\sum_{k=1}^{i} k P_{i,i-k} + \sum_{k=1}^{\infty} k P_{i,i+k} = D(i) + U(i) \quad (19)$$

where we have defined

$$D(i) = -\sum_{k=1}^{i} k \sum_{n=0}^{\infty} \lambda_n \sum_{j=k}^{i} B_i(j) \epsilon_{n+j,n+k}$$

$$U(i) = \sum_{k=1}^{\infty} k \sum_{n=0}^{\infty} \lambda_{n+k} \sum_{j=0}^{i} B_i(j) \epsilon_{j+k+n,n}. \quad (20)$$

The idea of the proof is to show that

$$\lim_{i\to\infty} [D'(i) + U'(i)] = -\theta \lim_{i\to\infty} d_i \quad (21)$$

and since it has been proved in Section III that $\lim_{i\to\infty} d_i = \lambda - C$, we will be able to conclude that $(X_n)_{n\geq 0}$ is transient for $\lambda > C$.

1) $\lim_{i\to\infty} [D'(i) + \theta D(i)] = 0.$

From (18) and (20)

$$D'(i)+\theta D(i)=(i+1)\sum_{k=1}^{i}\left[\left(\frac{i+1}{i+1-k}\right)^{\theta}-1-\frac{\theta k}{i+1}\right]\cdot\sum_{n=0}^{\infty}\lambda_n\sum_{j=k}^{i}B_i(j)\epsilon_{n+j,n+k}$$

which is more conveniently written as

$$D'(i)+\theta D(i)=(i+1)\sum_{n=0}^{\infty}\lambda_n\sum_{j=1}^{i}B_i(j)\sum_{k=1}^{j}\cdot\left[\left(\frac{i+1}{i+1-k}\right)^{\theta}-1-\frac{\theta k}{i+1}\right]\epsilon_{n+j,n+k}.$$

This expression is nonnegative since

$$\left(\frac{i+1}{i+1-k}\right)^{\theta}-1-\frac{\theta k}{i+1}>0\qquad(1\le k\le i).$$

Define $\gamma_k = \sup_{n\ge k}\epsilon_{nk}$. Then

$$0\le D'(i)+\theta D(i)\le(i+1)\sum_{n=0}^{\infty}\lambda_n\sum_{j=1}^{i}B_i(j)\sum_{k=1}^{j}\cdot\left[\left(\frac{i+1}{i+1-k}\right)^{\theta}-1-\frac{\theta k}{i+1}\right]\gamma_{n+k}$$

$$\le(i+1)\sum_{n=0}^{\infty}\lambda_n\sum_{k=1}^{i}\left[\left(\frac{i+1}{i+1-k}\right)^{\theta}-1-\frac{\theta k}{i+1}\right]\gamma_{n+k}.$$

That is

$$D'(i)+\theta D(i)\le x_1(i)+x_2(i)\tag{22}$$

with, assuming for instance that i is odd

$$x_1(i)=(i+1)\sum_{n=0}^{\infty}\lambda_n\sum_{k=1}^{(i+1)/2}\left[\left(\frac{i+1}{i+1-k}\right)^{\theta}-1-\frac{\theta k}{i+1}\right]\gamma_{n+k}$$

$$x_2(i)=(i+1)\sum_{n=0}^{\infty}\lambda_n\sum_{k=(i+3)/2}^{i}\left[\left(\frac{i+1}{i+1-k}\right)^{\theta}-1-\frac{\theta k}{i+1}\right]\gamma_{n+k}.\tag{23}$$

We show that $x_1(i)$ and $x_2(i)$ go to zero independently. Fix $\epsilon > 0$. Define for $0 < x \le i$ the function

$$p_i(x)=\frac{i+1}{x^2}\left[\left(\frac{i+1}{i+1-x}\right)^{\theta}-1\right]-\frac{\theta}{x}.$$

It is easily proved that for each $i \ge 1$, $p_i(x)$ is a positive nondecreasing function of x. Also

$$p\left(\frac{i+1}{2}\right)=\frac{1}{i+1}[4(2^{\theta}-1)-2\theta]=\frac{A}{i+1}$$

where A is a positive constant depending only on θ. From (23)

$$x_1(i)=\sum_{n=0}^{\infty}\lambda_n\sum_{k=1}^{(i+1)/2}k^2p_i(k)\gamma_{n+k}\le\frac{A}{i+1}\sum_{n=0}^{\infty}\lambda_n\sum_{k=1}^{n+(i+1)/2}k^2\gamma_{n+k}.$$

If $\lim_{k\to\infty}k^2\gamma_k = 0$, then $\lim_{n\to\infty}1/n\sum_{k=1}^{n}k^2\gamma_k = 0$. So we can choose i large enough so that for $n \ge (i+1)/2$, $\sum_{k=1}^{n}k^2\gamma_k < n\epsilon$. Then

$$x_1(i)\le\epsilon\frac{A}{i+1}\sum_{n=0}^{\infty}\lambda_n\left(n+\frac{i+1}{2}\right)=\epsilon A\left(\frac{\lambda}{i+1}+\frac{1}{2}\right).$$

Now if we choose i big enough so that for $k > (i+3)/2$, we have $\gamma_k < \epsilon/k^2$, then

$$x_2(i)\le\epsilon\sum_{n=0}^{\infty}\lambda_n\sum_{k=(i+3)/2}^{i}(i+1)\cdot\left[\left(\frac{i+1}{i+1-k}\right)^{\theta}-1-\frac{\theta k}{i+1}\right]\frac{1}{(n+k)^2}$$

$$\le\frac{4\epsilon}{i+3}\sum_{k=(i+3)/2}^{i}\left[\left(\frac{i+1}{i+1-k}\right)^{\theta}-1-\frac{\theta k}{i+1}\right].$$

By bounding the sum in the last equation by integrals, it can be seen that it is upper bounded by a linear function of i.

2) $\lim_{i\to\infty}[U'(i) + \theta U(i)] = 0$.

From (18) and (20)

$$U'(i)+\theta U(i)=\sum_{k=1}^{\infty}(i+1)\left[\left(\frac{i+1}{i+1+k}\right)^{\theta}-1+\frac{\theta k}{i+1}\right]\cdot\sum_{n=0}^{\infty}\lambda_{k+n}\sum_{j=0}^{i}B_i(j)\epsilon_{j+k+n,n}.$$

With a change of variable

$$U'(i)+\theta U(i)=(i+1)\sum_{j=0}^{i}B_i(j)\sum_{n=1}^{\infty}\lambda_n\sum_{k=1}^{n}\cdot\left[\left(\frac{i+1}{i+1+k}\right)^{\theta}-1+\frac{\theta k}{i+1}\right]\epsilon_{n+j,n-k}.$$

By using the following inequalities:

$$0\le\frac{1}{(1+x)^{\theta}}-1+\theta x\le\theta(1+\theta)\frac{x^2}{2}\qquad(x\ge0,\ 0<\theta<1)\}$$

we get

$$0\le U'(i)+\theta U(i)\le\theta\frac{(1+\theta)}{2}(i+1)\sum_{j=0}^{i}B_i(j)\cdot\sum_{n=1}^{N}\lambda_n\sum_{k=1}^{n}\frac{k^2}{(i+1)^2}\epsilon_{n+j,n-k}$$

$$+\theta(i+1)\sum_{j=0}^{i}B_i(j)\sum_{n=N+1}^{\infty}\lambda_n\sum_{k=1}^{n}\frac{k}{i+1}\epsilon_{n+j,n-k}$$

$$\le\frac{1}{i+1}\sum_{n=1}^{N}n^2\lambda_n+\sum_{n=N+1}^{\infty}n\lambda_n.$$

Fix $\epsilon > 0$. Choose N such that $\sum_{N+1}^{\infty}n\lambda_n < \epsilon/2$, and then, N being fixed, choose i large enough so that $1/(i+1)\sum_{n=1}^{N}n^2\lambda_n < \epsilon/2$. □

It should be clear at this point that unlike the ergodicity region, the recurrence region depends in general on the elements of the reception matrix (instead of only the row averages) and on the retransmission probability p. For this reason, an exact expression for the recurrence region seems rather difficult to obtain; nonetheless, the method (see [26]) that we used to study the example in Fig. 3 can be generalized to obtain the following upper and lower bounds on the recurrence region.

Theorem 4: $(X_n)_{n\ge 0}$ is recurrent for $\lambda < L$ and transient for $\lambda > U$, with $L = \max\{l_1, \sup_{0<\theta<1} l_\theta, \sup_{0<\theta<1} l'_\theta\}$ and $U = \min\{u_1, \inf_{0<\theta<1} u_\theta, \inf_{0<\theta<1} u'_\theta\}$ where

$$l_1 = \lim_{i\to\infty} (i+1) \sum_{n=0}^{\infty} \lambda_n \sum_{j=1}^{i} B_i(j) \sum_{k=1}^{n+j} \ln\left(\frac{i+n+1}{i+n-k+1}\right) \epsilon_{n+j,k}$$

$$l_\theta = \frac{1}{\theta} \lim_{i\to\infty} (i+1)^{1-\theta} \sum_{n=0}^{\infty} \lambda_n \sum_{j=1}^{i} B_i(j) \cdot \sum_{k=1}^{n+j} [(i+n+1)^\theta - (i+n-k+1)^\theta]\epsilon_{n+j,k}$$

$$l'_\theta = \frac{1}{\theta} \lim_{i\to\infty} (i+1)\,[\ln(i+1)]^{1-\theta} \sum_{n=0}^{\infty} \lambda_n \sum_{j=1}^{i} B_i(j) \cdot \sum_{k=1}^{n+j} [[\ln(i+n+1)]^\theta - [\ln(i+n-k+1)]^\theta]\epsilon_{n+j,k}$$

and

$$u_1 = \lim_{i\to\infty} (i+1)\,[\ln(i+1)]^2 \sum_{n=0}^{\infty} \lambda_n \sum_{j=1}^{i} B_i(j) \cdot \sum_{k=1}^{n+j} \left[\frac{1}{\ln(i+n+2-k)} - \frac{1}{\ln(i+n+2)}\right] \epsilon_{n+j,k}$$

$$u_\theta = \frac{1}{\theta} \lim_{i\to\infty} (i+1)^{1+\theta} \sum_{n=0}^{\infty} \lambda_n \sum_{j=1}^{i} B_i(j) \cdot \sum_{k=1}^{n+j} \left[\frac{1}{(i+n+1-k)^\theta} - \frac{1}{(i+n+1)^\theta}\right] \epsilon_{n+j,k}$$

$$u'_\theta = \frac{1}{\theta} (i+1)\,[\ln(i+1)]^{1+\theta} \sum_{n=0}^{\infty} \lambda_n \sum_{j=1}^{i} B_i(j) \cdot \sum_{k=1}^{n+j} \left[\frac{1}{[\ln(i+n+2-k)]^\theta} - \frac{1}{[\ln(i+n+2)]^\theta}\right] \epsilon_{n+j,k}.$$

We are assuming that the limits above exist, which indeed happens in most practical cases. The theorem is valid if any of these limits is infinite. In particular, if $L = +\infty$, then X_n is always recurrent. Note that usually, it is not necessary to carry out all the computations, because one of the three terms in the definition of L is equal to one of the terms in the definition of U. In fact, in most cases, we have $\sup_{0<\theta<1} l_\theta = \inf_{0<\theta<1} u_\theta$ if $0 < p < 1$, and $u_1 = l_1$ if $p = 1$. The proof of Theorem 4 can be found in [26].

References

[1] W. A. Rosenkrantz and D. Towsley, "On the instability of the slotted Aloha multiaccess algorithm," *IEEE Trans. Automat. Contr.*, vol. AC-28, pp. 994-996, 1983.

[2] G. Fayolle, E. Gelenbe, and J. Labetoulle, "Stability and optimal control of the packet switching broadcast channel," *J. Ass. Comput. Mach.*, vol. 24, no. 3, pp. 375-386, 1977.

[3] M. Kaplan, "A sufficient condition for nonergodicity of a Markov chain," *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 470-471, 1979.

[4] D. Bertsekas and R. Gallager, *Data Networks*. Englewood Cliffs, NJ: Prentice-Hall, 1987.

[5] B. Hajek and T. Van Loon, "Decentralized dynamic control of a multiaccess broadcast channel," *IEEE Trans. Automat. Contr.*, vol. AC-27, pp. 559-569, 1982.

[6] B. Hajek, "Hitting-time and occupation-time bounds implied by drift analysis with applications," *Adv. Appl. Prob.*, vol. 14, pp. 502-525, 1982.

[7] J. N. Tsitsiklis, "Analysis of a multiaccess control scheme," *IEEE Trans. Automat. Contr.*, vol. AC-32, pp. 1017-1020, 1987.

[8] S. Verdú, "Minimum probability of error for asynchronous Gaussian multiple-access channels," *IEEE Trans. Inform. Theory*, vol. IT-32, pp. 85-96, 1986.

[9] D. Raychaudhuri, "Performance analysis of random access packet-switched code division multiple access systems," *IEEE Trans. Commun.*, vol. COM-29, pp. 895-901, 1981.

[10] A. Polydoros and J. Silvester, "Slotted random access spread-spectrum networks: An analytical framework," *IEEE J. Select. Areas Commun.*, vol. SAC-5, no. 6, pp. 989-1002, 1987.

[11] M. K. Simon, J. K. Omura, R. A. Scholtz, and B. K. Levitt, *Spread-Spectrum Communications*. New York: Computer Science Press, 1985.

[12] N. Abramson, "The throughput of packet broadcasting channels," *IEEE Trans. Commun.*, vol. COM-25, pp. 117-128, 1977.

[13] D. J. Goodman and A. A. M. Saleh, "The near/far effect in local Aloha radio communications," *IEEE Trans. Vehicular Technol.*, vol. VT-36, Feb. 1987.

[14] L. G. Roberts, "Aloha packet system with and without slots and capture," *Comput. Commun. Rev.*, no. 5, pp. 28-42, 1975.

[15] C. Namislo, "Analysis of mobile radio slotted Aloha networks," *IEEE J. Select. Areas Commun.*, vol. SAC-2, pp. 583-588, 1984.

[16] D. H. Davis and S. A. Gronemeyer, "Performance of slotted Aloha random access with delay capture and randomized time of arrival," *IEEE Trans. Commun.*, vol. COM-28, pp. 703-710, 1980.

[17] J. C. Arnbak and W. van Blitterswijk, "Capacity of slotted Aloha in Rayleigh fading channels," *IEEE J. Select. Areas Commun.*, vol. SAC-5, no. 2, pp. 261-269, 1987.

[18] C. C. Lee, "Random signal levels for channel access in packet radio networks," *IEEE J. Select. Areas Commun.*, vol. SAC-5, no. 6, pp. 1026-1034, 1987.

[19] B. S. Tsybakov, V. A. Mikhailov, and N. B. Likhanov, "Bounds for packet transmission rate in a random multiple access system," *Problemy Peredachi Informatsii*, vol. 19, no. 1, pp. 61-81, 1983.

[20] N. Mehravari, "Collision resolution in random access systems with multiple reception," preprint, 1987.

[21] A. Shwartz and M. Sidi, "Erasure, capture and noise errors in controlled multiple access networks," in *Proc. 25th Conf. Decision Contr.*, Dec. 1986, pp. 1333-1334.

[22] R. E. Kahn, S. A. Gronemeyer, J. Burchfiel, and R. C. Kunzelman, "Advances in packet radio technology," *Proc. IEEE*, vol. 66, pp. 1468-1496, 1978.

[23] W. Szpankowski, "Analysis and stability considerations in a reservation multiaccess system," *IEEE Trans. Commun.*, vol. COM-31, no. 5, pp. 684-692, 1983.

[24] A. G. Pakes, "Some conditions for ergodicity and recurrence of Markov chains," *Oper. Res.*, vol 17, pp. 1058-1061, 1969.

[25] L. I. Sennott, P. A. Humblet, and R. L. Tweedie, "Mean drifts and the non-ergodicity of Markov chains," *Oper. Res.*, vol. 31, pp. 783-789, 1983.

[26] S. Ghez, S. Verdú, and S. C. Schwartz, "Random access communication with multipacket reception capability. I: Stability of Aloha," Dep. Elect. Eng., Princeton Univ., Tech. Rep. ISS. 8748, 1987.

[27] J. F. Mertens, E. Samuel-Cahn, and S. Zamir, "Necessary and sufficient conditions for recurrence and transience of Markov chains, in terms of inequalities," *J. Appl. Prob.*, vol. 15, pp. 848-851, 1978.

[28] C. M. Harris and P. G. Marlin, "A note on feedback queues with bulk service," *J. Assoc. Mach.*, vol. 19, pp. 727-733, 1972.

[29] F. G. Foster, "On the stochastic matrices associated with certain queuing processes," *Ann. Math. Statist.*, no. 24, pp. 355-360, 1953.

[30] P. Vit, "On the equivalence of certain Markov chains," *J. Appl. Prob.*, vol. 13, pp. 357-360, 1976.

[31] S. Ghez, S. Verdú, and S. C. Schwartz, "On decentralized control algorithms for multipacket Aloha," in *Proc. 25th Allerton Conf. on Commun., Contr., and Comput.*, Oct. 1987.

Sylvie Ghez (S'88) was born in Paris, France, on June 3, 1961. She received the engineering diploma from the Ecole Nationale Supérieure des Télécommunications, Paris, in 1984, and the M.S. degree in electrical engineering from Princeton University, Princeton, NJ, in 1987, where she is currently a doctoral candidate.

She is the recipient of the Wallace Memorial Fellowship for the year 1987–1988. Her research interests include multiple access protocols, network routing problems, queueing theory, and applied probability theory.

Sergio Verdú (S'80-M'84) received the Ph.D. degree in electrical engineering from the University of Illinois at Urbana-Champaign in 1984.

Upon completion of his doctorate he joined the faculty of Princeton University, Princeton, NJ, where he is an Assistant Professor of Electrical Engineering His recent research contributions are in the areas of multiuser communication and information theory, and statistical signal processing.

Dr. Verdú is a recipient of the National University Prize of Spain, the Rheinstein Outstanding Junior Faculty Award of the School of Engineering and Applied Science, Princeton University, and the NSF Presidential Young Investigator Award. He is currently serving as Associate Editor of the IEEE TRANSACTIONS ON AUTOMATIC CONTROL.

Stuart C. Schwartz (S'64-M'66-SM'83) was born in Brooklyn, NY, on July 12, 1939. He received the B.S. and M.S. degrees in aeronautical engineering from the Massachusetts Institute of Technology, Cambridge, in 1961, and the Ph.D. degree from the Information and Control Engineering Program, University of Michigan, Ann Arbor, in 1966.

While at M.I.T. he was associated with the Naval Supersonic Laboratory and the Instrumentation Laboratory. From 1961 to 1962 he was at the Jet Propulsion Laboratory, Pasadena, CA, working on problems in orbit estimation and telemetry. He is presently a Professor of Electrical Engineering and Chairman of the department at Princeton University, Princeton, NJ. He served as Associate Dean of the School of Engineering from July 1977 to June 1980. During the academic year 1972–1973, he was a John S. Guggenheim Fellow and a Visiting Associate Professor at the Department of Electrical Engineering, Technion—Israel Institute of Technology, Haifa, Israel. During the academic year 1980–1981, he was a member of the Technical Staff at the Radio Research Laboratory, Bell Telephone Laboratories, Crawford Hill, NJ. His principal research interests are in the application of probability and stochastic processes to problems in statistical communication and system theory.

Dr. Schwartz is a member of Sigma Gamma Tau, Eta Kappa Nu, and Sigma Xi. He has served as an Editor for the *SIAM Journal on Applied Mathematics* and as Program Chairman for the 1986 ISIT.

T-AC/34/9//29606

Control and Optimization Methods in Communication Network Problems

Anthony Ephremides
Sergio Verdú

Reprinted from
IEEE TRANSACTIONS ON AUTOMATIC CONTROL
Vol. 34, No. 9, September 1989

Control and Optimization Methods in Communication Network Problems

ANTHONY EPHREMIDES, FELLOW, IEEE, AND SERGIO VERDÚ, SENIOR MEMBER, IEEE

Abstract—In this paper we focus on two areas of communication network design in which methods of control and optimization theory have proven useful. These are the area of multiple access communication (for networks with shared links such as radio networks and local area networks) and the area of network routing (for networks with point-to-point interconnections). We review a few selected problems in each area to show the role of the control concepts involved and we then proceed to identify other areas of communication network design in which the same control theoretic and optimization methodology may be applicable and useful. We do not survey the work done in this area, nor do we review work in control areas whose methods are applicable in other communication network problems. Instead, we attempt to bring to the attention of the control systems community the numerous instances of problems arising in the pure communication network design process that can benefit from the attention and the capabilities of this community.

I. Introduction

Communication networks are designed and built in order to share resources. If interconnecting systems and bandwidths were available at no cost, then the solution to the problem of communication would be to assign dedicated communication links (channels) of sufficient capacity to every pair of conceivable users to meet their needs. This not being the case, it is necessary to multiplex the sources of communication traffic in order to optimize various cost criteria. Frequently, this optimization is dynamic and done on the basis of feedback that monitors the evolution of the degree of utilization of the network resources. Thus, we should expect a number of problems arising in communication network design to fit naturally in the framework of control systems design. In this paper we wish to demonstrate that indeed this is the case and to show how various control and optimization methodologies have been used in the study of communication networks.

In the beginning there was a single communication network, the telephone network. It represented a multibillion dollar investment and seemed to serve reasonably adequately the voice communication needs. The explosive growth in data communication needs during the last 30 years built up the pressure for additional and alternative networking options. As a result, the notion of store-and-forward switching (known also as message switching) was introduced in the early 1960's. This notion represented a breakthrough since it constituted a radical reversal of thinking with respect to the circuit-switching process; namely, instead of securing an open, dedicated "pipe" for the transmission of messages by means of hardware switches, it allowed a step-by-step (node-by-node) forwarding of messages, thereby permitting each node to switch messages by deciding when and where to transmit the messages in its buffer. In the last 20 years we have seen an avalanche of technologies (fast switching, time division multiplexing, local area networks, fiber optical networks, integrated services digital networks, etc.) and a proliferation of operational public and private networks that put these technologies to test and challenged communication engineers. In addition, they should challenge control engineers as well.

Without attempting a survey of this vast application area we wish to promulgate the viewpoint that many (if not most) specific sub-problems in the network design process are natural control problems. In support of this thesis, we choose, first, to demonstrate how two major areas in communication networks (routing and multiple access) have benefitted from the use of techniques borrowed from what is traditionally perceived as control systems methodology and, second, to mention additional areas that are likely to benefit from the control systems community. As illustrated in this paper, the techniques that have proved useful in communication networks include: dynamic programming (e.g., [2], [6], [8]–[10], [22], [29], [38], [39], [47], [49], [54]); linear programming (e.g., [50], [51]); constrained and iterative optimization (e.g., [5], [14], [16], [42]); Markov decision theory tools (e.g., [2], [26], [29], [38]); control of Markov chains (e.g., [11], [17], [18], [20], [40], [45]); stability analysis of stochastic systems via Lyapunov methods (e.g., [31], [43]); sample path dominance (e.g., [2], [52]); and convergence of distributed and asynchronous algorithms (e.g., [6], [16], [42]).

The problem of routing is encountered in all and every network that does not permit the source to reach the destination in a single transmission hop, but instead it must traverse a path of intermediate links. By contrast, the problem of multiple access is encountered primarily in those networks that permit the nodes to reach their destination directly in one hop by having to share the same link with other transmitting nodes. In addition, the two problems are fundamentally different in nature and, jointly, cover considerable ground in the networking area. Finally, together they facilitate the identification of additional design issues and the extension of the applicability of suitable control methods. Thus, they represent "cornerstone" areas of network design.

Routing can be studied either macroscopically or microscopically. The macroscopic viewpoint considers basically a flow model and determines the splitting of the flow in order to reach the destination in minimum time with efficient use of the network resources. It is traditionally referred to as *static routing*. The microscopic viewpoint dissects the flow process down to the atomic level of the individual transmission unit, the message (a string of bits commonly referred to as packet), and determines the path each message must follow at each of its hops through the network. It is traditionally referred to as *dynamic routing*. Both viewpoints are explored in Section II.

Multiple access is a collective term that refers to numerous problems that deal with the dynamic allocation of a single resource among users who can coordinate their use of that resource only by making use of that resource. These problems arise primarily in the context of radio channels but also in the

Manuscript received September 13, 1988; revised December 27, 1988 and April 24, 1989. Paper recommended by Associate Editor, T. L. Johnson. This work was supported in part by the National Science Foundation under Grants NSF-85-00108 and NSF ECS-88-57689, by the Office of Naval Research under Contracts N00014-84-K-0207 and N00014-87-K-0054, and by the Army Research Office under Contract DAAL-03-87-K-0062.

A. Ephremides is with the Department of Electrical Engineering, University of Maryland, College Park, MD 20742.

S. Verdú is with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544.

IEEE Log Number 8929606.

0018-9286/89/0900-0930$01.00 © 1989 IEEE

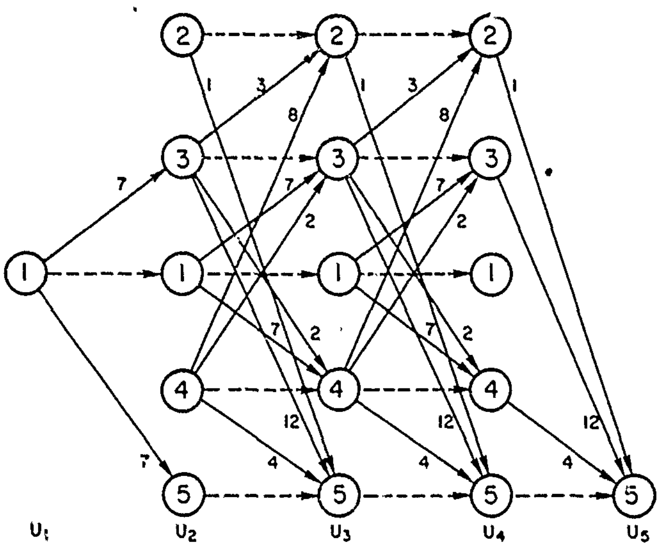


Fig. 1. Layered network showing link lengths. Source is node 1 in U_1 and destination is node 5 in U_5.

context of shared cable resources in local area networks. In Section III, we explore the main multiple access problems where control methods have been successfully applied.

Both in the case of routing as well as in the case of multiple access we place the emphasis on the control techniques that have been used. We then show how these techniques, sometimes with slight modification, can be naturally transported to other problem areas such as voice-data integration, flow control, and the scheduling of messages and links. This is done in Section IV.

II. Network Routing

The problem of routing in communication networks is one that has received early attention and has experienced significant breakthroughs in the brief history of the field of communication networks. It is one of the first problems that gained prominence as a result of the emergence of store-and-forward switching. It is also one in which analytical tools and available theories applied nicely from the beginning.

A. Static Routing

Given a network (a set of nodes connected by directed links) a path connecting the source node to the destination node has to be selected from the set of all possible such paths.[1] In the simplest formulation, the problem is one of finding the shortest path, i.e., a length is assigned to each link and the optimization criterion is the total path length. This problem is one of the archetypical combinatorial optimization problems (the solution can be found by exhaustive enumeration of a finite set of possibilities—all possible paths from source to destination). Among the many existing shortest path algorithms (see, e.g., [41]), the Bellman-Ford algorithm (1956) is of particular interest to our exposition, both because it is based on dynamic programming and because, as we will see below, it easily lends itself to distributed asynchronous implementation. A natural choice to find the shortest path from source to destination in a layered network (i.e., one in which the nodes can be grouped in subsets $U_1, \cdots U_M$ such that the source and destination nodes belong to U_1 and U_M, respectively, and there are links only between nodes in adjacent layers U_{k-1} and U_k) such as the one in Fig. 1, is the dynamic programming algorithm, where the shortest paths and distances (costs-to-go) of the nodes in layer U_k to the destination are computed based on the shortest paths and distances of the nodes in layer U_{k+1}. If the

[1] All the algorithms and results discussed in this section can be extended to the case where there are several source-destination pairs in the network.

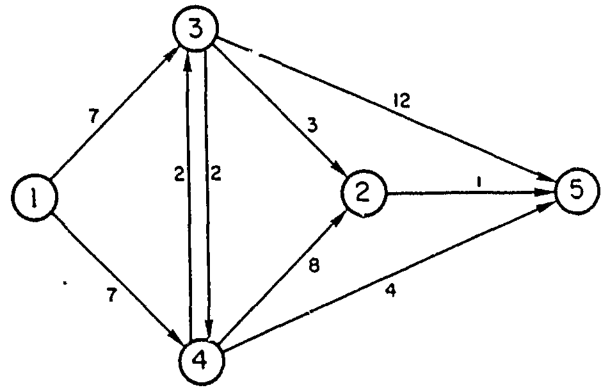


Fig. 2. Arbitrary network showing link lengths. Source is node 1 and destination is node 5.

network is not layered (such as that in Fig. 2), its shortest path can be obtained by finding the shortest path in a layered network derived from the original one as specified in the Bellman-Ford algorithm: the number of layers is equal to the number of nodes in the original network, say N, each layer contains a copy of each of the N nodes, and there is a link connecting two nodes in consecutive layers if such a link exists in the original network, in addition, copies of the same node in consecutive stages are connected by a zero-length link. (Fig. 1 was actually derived from Fig. 2 using this rule.) It is easy to see by induction that $D_k(i)$, the cost-to-go of node i in layer $N-k$, is the minimum length of any path from i to the destination that uses at most k links (in the original network). Since no shortest path uses more than $N-1$ links (link lengths are assumed nonnegative and, therefore, no path containing loops need be considered), the cost-to-go of node i at layer 1, $D_{N-1}(i)$ will indeed be the length of the shortest path from node i to the destination. Thus, the Bellman-Ford algorithm can be formulated as the iteration

$$D_k(i)=\min_{j\in N(i)}[D_{k-1}(j)+d_{ij}] \qquad \text{for } k=1,\cdots N-1 \quad (2.1)$$

where d_{ij} is the length of the link from i to j, $N(i)$ is the set of nodes for which such a link exists and it is assumed that $D_0(i)=\infty$ if i is not the destination node, which corresponds to the removal of all the nodes but the destination in the final layer (Fig. 1).

Contrary to what may appear at first glance there is a lot more to network routing than finding shortest paths. After all, the shortest path may not be the best path. The reason is that the real goal is to minimize the *delay* experienced in going from source to destination, and the delay encountered in each link is usually a function of the amount of traffic carried by the link (as the link becomes congested, it takes longer to go through it), which is referred to as the link flow and is quantified in packets (or messages) per second. Then, assuming a given desired flow level from source to destination, the problem is how to distribute it among all the possible paths so as to minimize the total delay. In contrast to the previous more elementary formulation of the routing problem which led to the shortest path combinatorial optimization problem and which corresponds to the special case in which the link delays are independent of the flows, we now face a continuous optimization problem which can be written as

$$\text{minimize } F(x)=\sum_{(i,j)} D_{ij}\left(\sum_{n\in P(i,j)} x(n)\right)$$

$$\text{subject to } x\in X=\left\{(x(1),\cdots x(J))\in R^J,\ \sum_{n=1}^{J} x(n)=\lambda,\ x(n)\geq 0\right\} \quad (2.2)$$

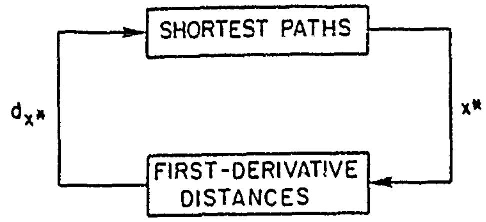


Fig. 3. Characterization of the solution to the minimum-delay routing problem.

where the set of all paths from source to destination is labeled $\{1, \cdots, J\}$; $x = (x(1), \cdots, x(J))$ is the vector of unknown nonnegative path flows which sum up to λ, the desired flow from source to destination; $P(i,j) \subset \{1, \cdots, J\}$ is the subset of paths that traverse link (i,j); and $D_{ij}(x)$ is the portion of the overall delay contributed by the link from node i to node j when the flow it carries is equal to x. In order to characterize a global solution to the optimization over a convex set in (2.2), it is natural to restrict attention to convex penalty functions. In practice, it is common that the incremental delay in a link grows with the amount of traffic it carries and, therefore, it can be assumed that the functions D_{ij} are convex without affecting significantly the practical applicability of the results.

Now, the characterization of the solution to (2.2), x^*, is straightforward. Since the feasible set X and the penalty function F are convex, it is necessary and sufficient that the directional derivative of the penalty function be nonnegative when evaluated at x^* in the direction of any of the elements of X (e.g., [37])

$$0 \le \lim_{\alpha \downarrow 0} \frac{1}{\alpha} [F((1-\alpha)x^* + \alpha x) - F(x^*)] \qquad \forall x \in X \quad (2.3)$$

which translates into

$$0 \le \sum_{(i,j)} D'_{ij}\left(\sum_{m \in P(i,j)} x^*(m)\right) \sum_{n \in P(i,j)} [x(n) - x^*(n)]$$
$$= \sum_{n=1}^{J} [x(n) - x^*(n)]\, d_{x^*}(n) \qquad \text{for all } x \in X \quad (2.4)$$

where $d_x(n) = \Sigma_{(i,j) \in L(n)} D'_{ij}(\Sigma_{m \in P(i,j)} x^*(m))$ is the length of path n when the length of each link is equal to the derivative of its delay evaluated at the set of flows x, and $L(n)$ is the set of links used by path n. The solution to (2.4), x^*, is the vector in X that minimizes its inner product with the vector of distances d_{x^*}. Thus, x^* puts all its weight on the smallest component(s) of d_{x^*}. The conclusion is that the optimum flow uses only shortest paths computed according to the *derivative* of the link delays.

This solution to the minimum-delay routing problem allows us to check whether a given set of flows is optimum. Unfortunately, it does not tell us how to find the optimum flows. Indeed, we face the chicken-and-egg situation depicted in Fig. 3. The optimum flows are obtained by solving a shortest path problem; but in order to compute the link lengths it is necessary to know the optimum flows. Nevertheless, the foregoing characterization of the optimal solution does suggest a possible iterative procedure to find the optimum set of flows. Starting with a given set of flows x one can compute the minimum derivative shortest paths for that flow, and hence, a new flow, $x^*(x)$ that is positive only along those shortest paths. The process can then be repeated, until there is no appreciable cost decrease. The region of convergence of such a procedure can be improved by letting the new flow be a convex combination of x and $x^*(x)$, i.e.,

$$x_{k+1} = (1-\alpha_k)x_k + \alpha_k x^*(x_k).$$

This is the so-called *flow deviation method* of Fratta, Gerla, and Kleinrock [14], where $0 \le \alpha_k \le 1$ is chosen to minimize

$$F((1-\alpha_k)x_k + \alpha_k x^*(x_k))$$

which is a special case of the feasible-direction nonlinear programming algorithm due to Frank and Wolfe [13]. The convergence of the flow deviation method to the optimum routing is rather slow because unfavorable paths tend to carry considerable flow during many iterations unless the initial routing guess is particularly fortuitous. Such a behavior can be improved by reducing the flow along each nonminimum derivative path in accordance to the delay experienced in that path. This is the idea of iterative routing algorithms based on *gradient projection* nonlinear optimization methods (e.g., [4]) in which the flow decrease along a nonminimum derivative path is proportional to the difference between its length and that of the shortest path (according to the first derivative of the delay function). If such a decrease would result in a negative flow, then the flow along that path is set to zero (hence, the *projection* to the set of feasible flows).

We have seen that the problem of static network routing can be formulated as a conceptually straightforward optimization problem that admits well-known solutions in nonlinear programming. What sets optimum routing in communication networks apart from other multicommodity flow problems arising in operations research is the fact that the optimization is carried out in real time, and often, in distributed fashion, where each node makes its own routing decisions based on local information. The review of centralized routing has revealed that the shortest path problem plays a central role in solving for the optimum routing regardless of whether the link congestion measures depend on the link flow or not. Hence, we will start the exposition of distributed routing algorithms by discussing the distributed version of the Bellman-Ford shortest path algorithm.

The Bellman-Ford updating equation in (2.1) suggests that the algorithm is suited for decentralized operation because each node can update its own estimate of distance to the destination (cost-to-go) provided it receives from its neighbors their own estimates [appearing on the right-hand side of (2.1)]. The feature that makes the study of the distributed Bellman-Ford algorithm interesting is that it can run completely *asynchronously*, in the sense that the updating and communication times need not be coordinated and convergence can be guaranteed by simply assuming that updating and communication between nodes never cease, without any requirements whatsoever on the rate of communication. The proof of convergence is a nice illustration of the analysis of decentralized algorithms where the processors are allowed to perform their computations and to communicate the corresponding results completely independently of one another [5], [6]. The idea is to show that the estimates computed in the distributed asynchronous algorithm are always sandwiched by the estimates computed by the centralized version of the algorithm when started at two different initial conditions, and that both centralized estimates converge to the true distances to the destination node.

Those centralized estimates are denoted by $\bar{D}_k = (\bar{D}_k(1), \cdots, \bar{D}_k(N))$ and $\underline{D}_k = (\underline{D}_k(1), \cdots, \underline{D}_k(N))$, and are the result of the centralized Bellman-Ford iteration (2.1) when it is started with initial conditions $\bar{D}_0 = (\infty, \cdots, \infty, 0)$ and $\underline{D}_0 = (0, \cdots, 0)$, respectively. (The destination node is assumed to be the Nth node.) Define the operator [see (2.1)]

$$B_i[D_k] = \min_{j \in N(i)} [D_k(j) + d_{ij}]$$
$$= D_{k+1}(i) \quad (2.6)$$

if $1 \le i < N$, and $B_N[D_k] = D_k(N)$. This operator is monotone in the sense that if $D \le D^*$ (i.e., if $D(i) \le D^*(i)$, $i = 1, \cdots N$), then

$$B_i[D] \le B_i[D^*]. \quad (2.7)$$

The monotonicity of B_i implies that

$$\underline{D}_k \le \underline{D}_{k+1} \le \bar{D}_{k+1} \le \bar{D}_k \quad (2.8)$$

and, moreover, it is easy to show that for sufficiently large k

$$\underline{D}_k = \bar{D}_{N-1} = \bar{D}_k \tag{2.9}$$

which is the vector of distances from each node to destination as we saw in the discussion of the centralized algorithm.

In the asynchronous distributed version of the algorithm, it is assumed that each node i keeps at time $t \geq 0$ an estimate of its distance to destination $A_t(i)$, and an estimate of the distance from each of its neighbors $j \in N(i)$ to destination $A_t^i(j)$, which is simply the latest estimate received from node j. In view of (2.8) and (2.9), convergence of the algorithm will follow if we show that for every index k, there exists a time $t_k > 0$ such that for all $t \geq t_k$

$$\underline{D}_k \leq A_t \leq \bar{D}_k \tag{2.10}$$

and for $i = 1, \cdots, N-1$

$$\underline{D}_k(j) \leq A_t^i(j) \leq \bar{D}_k(j) \qquad j \in N(i). \tag{2.11}$$

This is shown by induction. If $k = 0$, then (2.10) and (2.11) hold as long as the initial estimates of the decentralized algorithm are nonnegative. Assuming that the induction hypothesis is true for k, the monotonicity of B_i implies that if $t \geq t_k$, then

$$\underline{D}_{k+1}(i) = B_i[\underline{D}_k] \leq B_i[A_t^i] \leq B_i[\bar{D}_k] = \bar{D}_{k+1}(i). \tag{2.12}$$

But $A_t(i)$ is a piecewise constant function of time which only jumps at the updating times of node i, at which times it takes the value

$$A_t(i) = B_i[A_t^i].$$

Therefore, we can write

$$\underline{D}_{k+1}(i) \leq A_t(i) \leq \bar{D}_{k+1}(i) \qquad \text{for } t \geq t_k(i) \tag{2.13}$$

where $t_k(i)$ is the smallest updating time of node i which is greater than t_k. Moreover, if we wait long enough after $\max_i t_k(i)$, not only all the nodes will have carried out their first updates after t_k but the result of those computations will have been communicated to their neighbors because of the assumption that updating and communication occur infinitely often. Hence, there exists $t_{k+1} \geq \max_i t_k(i)$ such that for all $t \geq t_{k+1}$ and for all i and j

$$A_t^i(j) = A_s(j)$$

for some $s \geq t_k(j)$ (which depends on t, i, and j). Thus, it follows from (2.13) that

$$\underline{D}_{k+1}(j) \leq A_t^i(j) \leq \bar{D}_{k+1}(j) \qquad j \in N(i) \qquad i = 1, \cdots, N-1$$

completing the induction proof and, therefore, the proof of convergence of the distributed asynchronous Bellman-Ford algorithm.

When the link delays depend on the traffic flows, it is also possible to obtain the optimal routing that solves (2.2) in a distributed asynchronous fashion. Gradient projection algorithms are better suited for this task than the flow deviation method because in the latter method a higher degree of synchronization is required in order for the nodes to use the same step size at each iteration. In the distributed asynchronous implementation of gradient projection optimum routing algorithms, each node broadcasts from time to time the values of its outgoing flows to its upstream neighbors, who in turn pass that information on to their upstream neighbors. In this way, the source keeps estimates at all times of the link flows and can carry out the gradient projection iteration autonomously based on those estimates. The first algorithm based on this idea was due to Gallager [16], who posed an alternative formulation to (2.2), where the unknowns are the fractions of flow routed to each outgoing link at each node, rather

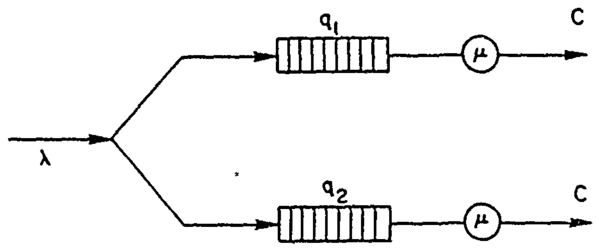


Fig. 4. Queueing model of a node with one incoming link and two outgoing links.

than the path flows. Tsitsiklis and Bertsekas [42] showed the convergence of the distributed asynchronous implementation of gradient projection optimal routing algorithms provided the time between consecutive broadcasts is small enough relative to the speed at which the flows generated by the algorithm change. The approach for showing the stability of this algorithm is very different from the proof of convergence of the distributed Bellman-Ford algorithm where the monotonicity of the dynamic programming mapping implied that the estimates are closer and closer to the solution regardless of the actual sequence of communication and computation times. The idea here is that if the step size of the algorithm is small enough, then the flows change so slowly with respect to the periods between communication times that their evolution is very close to that of the centralized algorithm which uses the unique, true value of each link flow.

B. Dynamic Routing

As mentioned earlier, there are two fundamentally different philosophies to network routing: either viewing it as a "flow" problem in which the traffic of messages is modeled as a "macro"-commodity entering the network as a single entity (static or quasi-static routing), or as an individualized-message path-finding problem in which the traffic is broken down to its constituent elementary units (dynamic routing)—a dichotomy akin to that of statistical/quantum mechanics in physics. Whereas the first approach leads to optimization problems where time plays no role, the essential ingredient of the second approach is the randomness of the time-evolution of the buffers in the network, thus placing dynamic routing within the sphere of stochastic control.

The most elementary instance of dynamic routing is the simple queueing system shown in Fig. 4 which models a node with one incoming link and two outgoing links. It simplifies considerably the dynamics of the message arrival process and of the service time characteristics and ignores processing delay. Thus, the arrival instants of messages over the incoming link are assumed to constitute a Poisson process of constant rate λ. Upon arrival each message is put in the buffer of one of the two outgoing links. This action represents the "control." The buffers are assumed to have unlimited (infinite) capacity and the message lengths are assumed to be random with exponential distribution (an obvious additional simplification) with parameter μ. The two outgoing links have equal capacity of C bits/s. Thus, each link is modeled as a queueing system with exponential service time distribution with parameter μC. It is desired to characterize the optimal control policy that minimizes the average total delay per message based on the observations of the "state" of the system, namely the number of messages q_1 and q_2 in the two buffers. The model, of course, assumes that the head-of-the-line message is dropped from the buffer as soon as the transmission of its last bit is completed.

This model, despite its simplicity, proved to be rather difficult to analyze. For details, see [10]; it is not important to repeat them here. It should suffice to state that the main result, which simply requires that upon arrival a message should join the shortest queue (with arbitrary decision in case the two queues have equal numbers of messages), was hardly surprising. Yet an intricate argument on the dynamic programming equation (DPE) was needed and there were some counter-intuitive side-results including the relaxation of the Poisson assumption on the arrivals, and the fact that in the incomplete state information case, the *certainty-equivalent control* (i.e., send the message to the *expected* shortest queue) need not be optimum unless both queues have the same number of customers initially.

The optimality of the send-to-shortest-queue (SS) policy in the complete state information case can be proved in a rather strong sense. At all times, the sum $(q_1 + q_2)$ and maximum $(\max\{q_1, q_2\})$ of the number of messages in both buffers are stochastically minimized by the SS policy in the sense of the partial order between random variables according to which the random variable X is *stochastically smaller* than Y if $P[X \le a] \ge P[Y \le a]$ for all a. The proof of optimality can be obtained by the method of *forward induction* [53], whereby the desired stochastic ordering between the queue sizes under the optimum and an arbitrary policy is shown to be preserved at each transition.

The problem formulation of [10] is one of many related ones (see [8], [9], [22], [24], [33], [38], [54], [55]) which are slightly more complicated but share some fundamental characteristics which, in fact, extend beyond the confines of the routing problem into the areas of priority assignment, resource allocation, and flow control. They are all Markovian decision process (MDP) problems. In the sequel we will describe a fairly general MDP that includes the dynamic routing problem as a special case. In fact, it includes almost all of the queueing control problems that have been studied in connection with communication network issues. We will then outline the solution methodologies that have been used. These include basically: 1) the derivation of optimality conditions from the DPE associated with the corresponding MDP; 2) the use of sample path stochastic dominance arguments, and finally; 3) the reformulation of the MDP as a linear program. We should emphasize, lest the reader be unduly encouraged, that the problems in this area are sufficiently complex, so that only modest results can be generally obtained despite involved arguments and nontrivial machinery. Typically, these results characterize some structural properties of the optimal policy. However, knowledge of such structure is often sufficient to permit close approximation of the actual optimal policy by well-founded heuristics.

Let us recall briefly what an MDP is (for details, see [30]). We need a state description of the process to be controlled. Let S be its state space. When in state $s \in S$, a set A_s of admissible control actions is specified. When action $a \in A_s$ is applied, there is a transition from state s to s' that is governed by the probability distribution $p(s'|s, a)$, and which occurs after a random time τ which is exponentially distributed with distribution denoted by $t(\tau|s, a, s')$. Clearly, p and t together describe the stochastic dynamics of the process to be controlled. Finally, each transition is accompanied by a cost penalty that we denote by $c(\tau, s, a, s')$.

The dynamic routing problem we considered before fits in this formulation easily. In that case, the state space is $S = \{0, 1, 2, 3, \cdots\}^2$. An element $s = (q_1, q_2) \in S$ is simply the pair of values of the respective queue sizes. The set of actions A_s is the same for any state and consists of a_1 and a_2 where a_i is the action that assigns an arriving message to the buffer of link i. The distribution p is of trivial form, in that the transitions are deterministic. Assignment of an arrival to queue i augments q_i by one. Note, now, that in addition to the arrival instants, the departure (or service completion) instants are important because they induce state transitions as well. A departure from queue i reduces q_i by one. When a departure occurs there is no meaningful control action that can be applied in this particular problem. The exponential distribution t corresponds to times between arrivals and/or departures.[2] Finally, the cost rate c must reflect the delay. By Little's result in queueing theory, we know that the average delay is proportional to the average number of customers in the queue. Thus, $c(\tau, a, s, s')$ can be taken to be simply equal to $(q_1 + q_2)$. This MDP formulation can be extended to encompass more complicated queueing control problems.

Let us return now to the general MDP. We need to specify the notion of a control policy and the optimization criterion. Let us denote by $\xi_1, \xi_2, \cdots$, the state transitions that occur at instants $t_1, t_2, \cdots$. A policy π is a sequence of decision rules $\pi_1, \pi_2, \cdots$, where π_n determines the choice of action at the transition time t_n. It can be viewed as a conditional distribution on the set of actions parametrized by the past history of the process.

The optimization criterion that corresponds to the practical case of expected total delay is the long-run average expected cost; namely, if we denote by $V(\pi, i, t)$ the expected cost incurred under policy π, with initial state i, until time t we consider as the optimization criterion the value function

$$V(\pi, i) \triangleq \liminf_{t\to\infty} \frac{V(\pi, i, t)}{t}.$$

For technical reasons, however, that are well known to optimization specialists, it is easier to establish optimality conditions if we consider, instead, the so-called α-discounted cost, i.e.,

$$V^\alpha(\pi, i) = \int_{t=0}^{\infty} e^{-\alpha t}\, dV(\pi, i, t).$$

The latter converges to the former as $\alpha \to 1$ under a variety of stationarity conditions. For technical reasons that will become apparent in the sequel, we will also consider the finite-horizon costs. These are defined in a similar fashion except that we let time extend only to t_n, the instant of the nth transition. If we denote by $V^\alpha(i)$ and $V(i)$ (and also $V_n^\alpha(i)$, $V_n(i)$ for the finite horizon cases) the values of these cost functions when π is chosen optimally, we are led to the following DPE:

$$V^\alpha(i) = \inf_{\alpha\in A_i} \sum_{i'} [c(i, a, i') + \beta(i, a, i')V^\alpha(i')]p(i'|a, i)$$

or

$$V_{n+1}^\alpha(i) = \inf_{\alpha\in A_i} \sum_{i'} [c(i, a, i') + \beta(i, a, i')V_n^\alpha(i')]p(i'|a, i)$$

where

$$\beta(s, a, s') \triangleq \int_0^\infty e^{-\alpha\tau}\, dt(\tau|s, a, s')$$

and

$$c(s, a, s') \triangleq \int_0^\infty c(\tau, s, a, s')\, dt\,(\tau|s, a, s')$$

are the discount factor and cost values per transition, respectively.

The DPE is of fundamental importance in the study of MDP's because the value function V^α has the usually convenient properties of convexity, supermodularity, and other forms of monotonicity that lead readily to sufficient conditions for optimality. The difficulty with the analysis of the DPE is that the optimality conditions are heavily problem-dependent and often lead to explosively large numbers of cases to be verified separately. This is especially true for MDP's that arise from queueing models. For this reason, and because of additional difficulties that arise when the state is on the boundaries (see [22]), it became evident that alternative methods of solution were needed.

[2] A slight modification of the model of transitions, called uniformization, is useful in that it introduces dummy transitions from a state into itself: thus, some situations which introduce nonessential complications can be handled without departure from this discrete transition time formulation.

One alternative method that has received attention recently and which produced successful results in problems of queueing control (akin to the routing problem) is a probabilistic method called sample-path or stochastic dominance. This method bypasses completely dealing with the value function. Instead, it focuses directly on seeking the optimal policy. Let G be the class of admissible policies. If we suspect that the optimal policy π has a property p, then we can proceed as follows in order to prove that it actually does have that property. Let S be a subset of G, to which we know the optimal policy belongs. We consider a subset of policies $S_p \subset S$, all elements of which have the property p. For every $\pi \notin S_p$, we attempt to construct a policy $\tilde{\pi}$ which outperforms π. If we succeed, we must conclude that the optimal policy belongs to S_p. In constructing $\tilde{\pi}$ we often need to engage in a careful reorganization of the underlying probability space in order to align the sample paths properly, so that the comparison of the two policies can be made for every sample path. This procedure is full of risks and extreme care is required to avoid faulty arguments. Note, also, that to apply this method usefully, we must have "guessed" the properties of the optimal policy correctly. Thus, at best, it is a method to verify the validity of our conclusions, rather than a method that leads us to the right conclusions.

Successful use of the stochastic dominance approach was made in [52] and [50] where a problem that is dual to the problem of dynamic routing was studied. Specifically, in a two-server queueing system in which the two servers have unequal service rates, we wish to determine whether and when the slower server needs to be activated if we are interested in minimizing the usual total expected delay function. That the optimal policy has a threshold form (namely that the slower server must be activated when the queue size exceeds a crucial value) was proven in [29] via the DPE method. However, the alternative proof via the arguments of stochastic dominance was much simpler and led to a generalization of the result to cases of nonexponential arrivals and/or service, that could not have been easily accomplished by means of the DPE method.

Another successful use of the stochastic dominance method has been noted in [2]. In this case the problem of optimally choosing which customer to serve next in a single queueing system was considered under the constraint that each customer must begin (or terminate) service by an individually assigned random deadline or else it is dropped from the system. The cost criterion is then to minimize the expected number of lost customers. It was proven that scheduling the customer with shortest time to extinction minimizes this cost.

Although these problems differ from routing, the model structures are quite similar, and it has been observed that, usually, queueing control problems with such structural similarities can be studied equally successfully.

The third method, which was first used in [38] in the study of a specific queueing control problem, and which has been broadly extended recently in [51], is the linear programming approach. Almost any queueing control problem that can be formulated as a MDP (therefore the problem of dynamic routing, as well) can be converted to an equivalent linear program (LP). The advantages of this conversion are that it is problem-independent and it leads occasionally to successful study of semi-Markov decision problems as well. Furthermore, it facilitates considerably the characterization of optimal solution properties. Here is how this equivalence can be demonstrated.

Let us concentrate on an MDP under a finite-horizon, discounted cost formulation.[3] We shall consider a queueing model with state dynamics given by

$$x_{k+1} = x_k + \xi_{k+1} z_{k+1}.$$

Here, x_k denotes the state at t_k (the instant of the kth transition), ξ_k represents that transition, and z_k represents the control action at that transition. The transition ξ_k can represent an arrival or a departure as an increment of the state. The control z_k is conveniently defined to enable ($z_k = 1$) or disable ($z_k = 0$) a transition. For example, in the routing model discussed at the beginning of the section, the state is equal to a two-dimensional vector of queue sizes, and the transition corresponding to sending an arriving message to the first queue would be represented by $\xi_k = [1\ 0]^T$. Indeed, a variety of queueing control problems (in fact, the vast majority of those that have been considered in connection with communication network problems) can be so represented.

Note that the crucial aspect of this state equation is the *linear* dependence on the controls. Note also that usually the cost function is linear in the state (since the usual cost criterion is the expected delay which is coupled to the queue sizes, and hence the state, by Little's result). Consequently, the cost is linear in the controls. The minimization of the cost over the set of control trajectories is constrained since the state equation must be satisfied and the state must always belong to an admissible set (typically, a set of vectors with integer-valued coordinates belonging to given ranges). Thus, the constraints are also linear in the controls, and the problem is easily formulated as an LP. There are, however, two points that require attention. First, the controls are integer-valued, i.e., $z_k \in \{0, 1\}$. Second, in the MDP the vectors ξ_k are random and depend on past history.

The first problem is taken care of in one of two ways: by construction or by use of a property of the constraint matrix of the linear program, called unimodularity. The construction method involves using a noninteger optimum control whose quantized version satisfies the MDP optimality conditions (see [38], [51] for details). The use of unimodularity involves a well-known result in the theory of integer linear programming (e.g., [34]): if the constraint matrix of an LP is integer-valued and totally unimodular (i.e., each of its sub-determinants is $+1$, -1, or 0), then all the vertices of the feasible polytope are integer-valued. Therefore, no further restrictions are needed to guarantee that the solution of a conventional LP will result in the integer-valued optimal control. Fortunately, in many queueing problems of interest (including the dynamic routing problem), the constraint matrix is indeed totally unimodular.

The second problem is easily taken care of by thinking of the z_k's as functions from the sample space Ω to the action space. Thus, the cost criterion can be written as a functional on the underlying probability space.

Let $z_k(\omega_k)$ represent the control action at the kth transition, where ω_k denotes the random "history" until the kth transition. We have

$$x_{k+1}(\omega_{k+1}) = x_k(\omega_k) + z_{k+1}(\omega_{k+1})\xi_{k+1}(\omega_{k+1}).$$

Let S and Z be the set of admissible states and controls, respectively. The β-discounted, n-step, expected cost under policy z and initial condition x is given by

$$J_n^\beta(x, z) = E_z \sum_{k=0}^{n-1} \beta^k L(z_k)$$

where

$$L(z_k) = c^T x_k + d^T z_k$$

(c and d denote constant column vectors). This is a cost function that is adequately general. For example, in a pure resource allocation problem without blocking or rejection of messages we have $d = 0$, while in pure blocking problems we take $c = 0$. The state equation, after repeated iterations, yields

$$x_k(\omega_k) = x + \sum_{j=1}^{k} z_j(\omega_j)\xi_j(\omega_j), \qquad k > 0.$$

[3] The reason that we cannot work directly with infinite horizons is the possibility of so-called duality gaps in linear programming theory with infinite-dimensional variables.

Therefore,

$$J_n^\beta(x, z) = E_x \sum_{k=0}^{n-1} \beta^k \left\{ c^T x + c^T \sum_{j=1}^{k} z_j \xi_j + d^T z_k \right\}$$

$$= \frac{1-\beta^n}{1-\beta} c^T x + E_x \sum_{k=1}^{n} \beta^k \left\{ \sum_{j=1}^{k} c^T z_j \xi_j + d^T z_k \right\}.$$

But

$$E_x(z_k) = \sum_{\omega_k} z_k(\omega_k) Pr(\omega_k).$$

Hence

$$J_n^\beta(x, z) = \frac{1-\beta^n}{1-\beta} c^T x + \sum_{k=1}^{n} \sum_{\omega_k} \gamma_k(\omega_k) z_k(\omega_k)$$

where $\gamma_k(\omega_k)$ is a known function that depends on $Pr(\omega_k)$, c, ξ_k, and β^k. Consequently, the MDP is equivalent to

$$\min_{z_k} \sum_{k=1}^{n} \sum_{\omega_k} \gamma_k(\omega_k) z_k(\omega_k)$$

subject to

$$\left(x + \sum_{j=1}^{k} z_j(\omega_j) \xi_j(\omega_j) \right) \in S$$

which is a conventional LP where the initial condition plays the role of a parameter, the sensitivity with respect to which can be studied by the well-developed theory of sensitivity analysis of linear programming [15].

In conclusion, we see that the MDP is converted to an equivalent LP under very mild conditions that are usually satisfied by dynamic routing and other queueing control problems. Thus, a third alternative methodology becomes generally available for the study of these problems. Whether to choose from the arsenal the DPE approach, or the LP method, or stochastic dominance tools, depends on the problem and on the, as yet undeveloped, intuition that the investigator should possess.

III. Multiple-Access Communications

The communication networks considered in the discussion of routing problems in Section II consist typically of a set of nodes connected by point-to-point communication links. Each of these links viewed in isolation can be modeled as a classical communication channel with one sender and one receiver. In this section, we consider multipoint-to-point communication links where several transmitters share a common channel. Multiple-access channels are the basic building blocks of radio networks, satellite communication, and local area networks, and during the last 15 years have attracted the attention of many communication. information, and control theorists.

There is a wide variety of strategies to divide the "resources" of a communication channel among several geographically dispersed transmitters. The simplest methods are those that assign a permanent independent sub-channel to each transmitter (e.g., in frequency division multiple access and time division multiple access); these strategies are easy to analyze and are widely used in practice in situations where the users need to transmit at fairly steady rates. If the transmitters are *bursty* (i.e., the radio of peak-to-average rate at which the need to transmit is high) those static methods are inefficient since most of the time the channel is underutilized while demand (and induced delay) accumulates at busy terminal locations. Dynamic channel sharing strategies overcome this problem by allocating channel resources on an on-demand basis. Consistent with the overall spirit of this paper, our goal here is not to review this vast topic, but rather to demonstrate how control theory can play a useful role in its study. Here we wish to single out two multiple access strategies: random access and simultaneous transmission, which are broadly representative of dynamic channel sharing systems and in which control theoretic concepts have played a pivotal role.

In random access communication, the conceptual allocation model is addressed without an effort to exploit the signaling degrees of freedom and the micro-structure of the transmitted messages. For this purpose, a crude channel model is considered, that achieves this separation of the "macro" from the "micro" problem. In simultaneous transmission systems, however, a more refined viewpoint is adopted, by taking the realities of the medium into account, modeling them, and exploiting them.

A. Random-Access

The object of interest here is the so-called *collision channel* model, in which messages (called packets) require one time unit (called slot) for transmission and are sent by a population of users who are synchronized so that their slots coincide at the receiver, but are otherwise uncoordinated and unaware of which and how many users have packets to transmit. If two or more packets are simultaneously transmitted, it is assumed that the receiver is unable to recover any of the messages, and they have to be retransmitted in a future slot. In the ALOHA algorithm, which was developed in the early 1970's [1] at the University of Hawaii and marked the beginning of the area of random-access communication, each packet that has been unsuccessfully transmitted before is transmitted with probability p in the next slot. New packets which have not attempted transmission before are transmitted with probability either 1 or p depending on which version of the ALOHA algorithm is used. In our discussion, we will assume the latter choice.

Under these conditions, and assuming that the number of newly generated packets in each slot is a random variable (with mean λ) independent from slot to slot, the number of packets awaiting transmission (called backlog) is a Markov chain taking values in $\{0, 1, 2, \cdots\}$. The central problem is to investigate under what conditions the backlog Markov chain is *ergodic*, i.e., it is stable in the sense that it reaches a steady state in which the periods between the times when there are no packets to transmit are not too infrequent (they have finite expected value). The transition probabilities of the Markov chain are parametrized by the rate of arrival of new packets λ and the retransmission probability p. Whereas λ is fixed and given, p is chosen by the transmitters. Hence, we are dealing with a fairly simple *controlled Markov chain* whose control space is the interval $(0, 1]$. In the original ALOHA algorithm, the control p remained constant and common to all transmitters regardless of the information acquired by listening to the channel, thereby resulting in the open-loop control of the Markov chain. Depite several "proofs" of the stabilitiy of ALOHA published during the 1970's, neither the actual system built in Hawaii nor the ideal Markov chain model were stable. The reason why the open-loop system is unstable can be easily understood by considering the backlog drift, $d(n)$, which is defined as the expected increase in the backlog over the next slot when the current value of the backlog is equal to n. It is easy to see that the backlog drift is given simply by the expected number of new packets per slot minus the expected number of successfully transmitted packets in the next slot, i.e.,

$$d(n) = \lambda - [np(1-p)^{n-1}]. \tag{3.1}$$

The drift quantifies the expected evolution of the Markov chain from each state, and therefore it is a valuable tool in analyzing the stability of the chain. For any $p \in (0, 1]$ the term in brackets in

(3.1) goes to 0 as $n \to \infty$, and hence, the drift is positive and close to λ for sufficiently large backlogs. This implies that when the backlog is large it tends to grow, thereby eliminating any hope for stability. Using standard results, this reasoning can be formalized straightforwardly to prove not only the instability of the open-loop system [11] for all values of λ and p, but the fact that the backlog goes to infinity with probability one [25], [35], [40].

Fortunately, the system can be stabilized by closed-loop control. Let us examine first the case of complete-state information, i.e., each station is informed at the end of each slot of the current value of the backlog and chooses the retransmission probability on the basis of that information. As far as stability is concerned, the best choice of the retransmission probability p is the value that minimizes the drift because that results in the maximum possible arrival rate that guarantees stability (called the throughput). It follows from (3.1) that the optimum value of p is

$$p^*(n)=\frac{1}{n}, \qquad n=1, 2, \cdots \tag{3.2}$$

and the resulting drift is

$$d^*(n)=\lambda-\left[1-\frac{1}{n}\right]^{n-1} \tag{3.3}$$

which is negative for $n > 1$ when $\lambda < e^{-1}$, and is positive for large backlogs when $\lambda > e^{-1}$. Therefore, the throughput of the closed-loop system with complete state information is $e^{-1} = 0.368$. However, the relevance of complete state information feedback is rather limited in practice. This is because the instantaneous value of the backlog is available to each station only if there exists so large a degree of communication among the transmitters that much more efficient algorithms than ALOHA can be used.

The case of partial state information is the problem of interest in practice, since the only feedback available to each station is the outcome (collision, success, empty) of the transmission in each slot. The analysis of the controlled system with partial state information was pioneered by Hajek and Van Loon [20] who proposed a recursive updating law of the retransmission probabilities as a function of the channel outcomes. This feedback policy was shown in [21] to attain the throughput achievable with complete-state information, namely e^{-1}. Those papers and subsequent works have referred to the problem as *decentralized control* of ALOHA, motivated by the fact that each station chooses the retransmission probability autonomously based on the channel feedb[illegible] [illegible]owever, it is useful to recognize that the problem boi[illegible] to (centralized) stochastic control with one decision m[illegible] and incomplete state information because all stations are constrained to use the same retransmission probabilities.

We will review here the proof of stability of the following *certainty-equivalence* closed-loop control:

$$p(\hat{n})=\frac{1}{\hat{n}} \tag{3.4}$$

where $\hat{n}$ is an estimate of the backlog updated according to

$$\hat{n}_{k+1}=\begin{cases}\max\{1, \hat{n}_k+\alpha\} & k\text{th slot is idle}\\ \hat{n}_k+\beta & k\text{th slot is success or collision.}\end{cases} \tag{3.5}$$

The throughput attainable with this feedback law depends on the constants $\alpha < 0$ and $\beta > 0$. As we will see, there exists a set of choices for those constants that results in throughput equal to e^{-1}.

Unlike the case of complete-state information, the proof of stability is not straightforward because now it is the two-dimensional process formed by the backlog and its estimate $\{(n_k, \hat{n}_k)\}_k$ (rather than the backlog itself) which is a Markov process. According to (3.4) and (3.5) the drift of this Markov process is given by

$$\begin{aligned}&E[(n_{k+1}, \hat{n}_{k+1})-(n_k, \hat{n}_k)|(n_k, \hat{n}_k)=(n, s)]\\ &=\left(\lambda-\frac{n}{s}\left[1-\frac{1}{s}\right]^{n-1}, \beta+(\max\{\alpha, 1-s\}-\beta)\left[1-\frac{1}{s}\right]^{n}\right)\\ &\triangleq (d(n, s), c(n, s)).\end{aligned} \tag{3.6}$$

Contrary to what we saw in the case when the state is known, it is not true that the backlog drift is negative for sufficiently large backlogs. As we can see in Fig. 5, if the estimate is far from the true value, then the backlog may actually tend to increase.

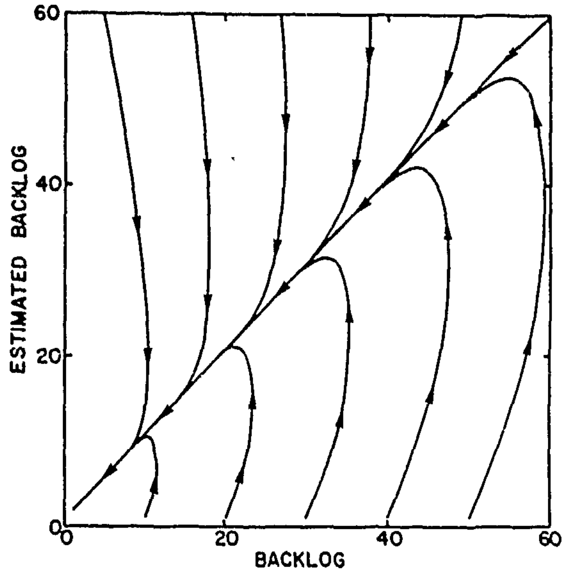


Fig. 5. Drift of (backlog, backlog estimate) Markov process for decentralized control with $\alpha = -1.48$, $\beta = 0.8$, and $\lambda = 0.33$.

However, at every point in the state space the tendency of the process is to approach the diagonal where the estimate is equal to the true value of the backlog. Furthermore, as Fig. 5 or the analysis of the perfect-state information case shows, the drift along the diagonal is negative. Such a behavior is a strong indication of the stability of the controlled Markov process.

This can be proved using a powerful sufficient condition found by Mikhailov [31] for the stability of a Markov process taking values in $\mathbb{R}^+ \times \mathbb{R}^+$. In essence, Mikhailov's condition states that it is enough to restrict attention to those points of the state space where either the backlog or its estimate are large and at which the drift is *radial*, i.e.,

$$\frac{d(n, s)}{c(n, s)}=\frac{n}{s};$$

then, it is sufficient for stability that the drift point towards the origin at those states. To see that this condition is indeed satisfied for our system, we compute first the asymptotic drifts along the radius $\{(n, s): n/s = \psi\}$ for $\psi \in [0, \infty)$

$$d(\psi)=\lim_{s\to\infty} d(\psi s, s)=\lambda-\psi e^{-\psi} \tag{3.7a}$$

$$c(\psi)=\lim_{s\to\infty} c(\psi s, s)=\beta+(\alpha-\beta)e^{-\psi}. \tag{3.7b}$$

It can be checked using (3.7) that if the constants α and β in (3.5) are chosen such that $\beta > 0.23\lambda$ and $\lambda - e^{-1} = \beta + (\alpha - \beta)e^{-1}$, then the drift is radial only at $\psi = 1$ (cf. Fig. 5), where it points towards the origin as long as $d(1) = \lambda - e^{-1} < 0$.

Mikhailov's sufficient condition can be justified constructing a stochastic Lyapunov function to prove the stability of a Markov process $\{x_k\}_k$ with state space $\mathbb{R}^+ \times \mathbb{R}^+$. To that end, it is advantageous to switch to polar coordinates (r, ϕ) and to define the radial drift $\delta(r, \phi)$ as the projection of the drift along the direction of the point (r, ϕ) and the tangential drift $\mu(r, \phi)$ as the projection of the drift along the direction perpendicular to (r, ϕ). Denote the asymptotic drifts $\delta(\phi) = \lim_{r\to\infty} \delta(r, \phi)$ and $\mu(\phi) = \lim_{r\to\infty} \mu(r, \phi)$ and define the function

$$V(r, \phi) = r\Phi(\phi)$$

where

$$\Phi(\phi) = \exp\left(-C \int_0^\phi \mu(v)\, dv\right) \qquad \phi \in \left[0, \frac{\pi}{2}\right].$$

Note that $V(r, \phi)$ is a candidate Lyapunov function because it is positive outside the origin and $V(r, \phi) \to \infty$ as $r \to \infty$. Furthermore, it can be shown [31] that the asymptotic drift of the candidate Lyapunov function is equal to

$$\lim_{r\to\infty} E[V(x_{k+1}) - V(x_k) | x_k = (r, \phi)] = \Phi(\phi)[\delta(\phi) - C\mu^2(\phi)]. \quad (3.8)$$

Now, under Mikhailov's condition, the asymptotic drifts are assumed continuous on $[0, \pi/2]$ and $\delta(\phi) < \epsilon$ for any phase such that $\mu(\phi) = 0$ (i.e., whenever the drift is radial it points towards the origin), therefore, the constant C can be chosen large enough so that the left side of (3.8) is upper bounded by a negative constant. This implies that $V(r, \phi)$ is indeed a stochastic Lyapunov function and therefore standard results on the stability of stochastic systems [27], [45] can be applied to show the stability of the system.[4]

In some multiaccess environments, the receiver can indeed demodulate reliably one or more packets even in the presence of other interfering packets and the collision channel model no longer applies to those cases. The results reviewed in this section can be generalized to a general channel with *multipacket reception* capability, to show that: 1) the throughput of open-loop ALOHA is equal to the limit of the expected number of successfully received packets per slot as the backlog goes to infinity [17]; and 2) the throughput of closed-loop ALOHA (with either complete or partial state information) is equal to the maximum over v of the expected number of successfully received packets per slot when the number of attempted transmissions is a Poisson random variable with mean v [18].

Returning to the case of the collision channel, the next natural step is to drop the main restriction in the ALOHA algorithm, namely, that all stations use the same retransmission probability. This is done in a class of random-access algorithms referred to as collision resolution algorithms which are characterized by the fact that not only are all blocked packets eventually retransmitted successfully, but all users eventually become aware that these packets have been successfully retransmitted. Contrary to the ALOHA algorithm, the decision whether or not to transmit a packet takes into account the previous history of attempted retransmissions of that particular packet. The introduction of this new dimension into the problem renders Markov chain tools considerably less useful than in the foregoing analysis and converts it into a very difficult decentralized stochastic control problem, for which the optimum throughput remains unknown[5] despite many efforts.

[4] Another choice of stochastic Lyapunov function for the specific case of decentralized control of ALOHA can be found in [43].

[5] The best known algorithm has been shown to achieve a throughput of 0.488 using Howard's policy iteration for sequential infinite-horizon problems [32] or by reduction to a simple optimization problem [48]. On the other hand, it is known that the optimum throughput is upper bounded by 0.568 [44].

B. Simultaneous Transmission

In contrast to random-access communication systems, in simultaneous transmission multiple-access systems, the transmitters send their messages simultaneously, independently, and without monitoring the channel in any way. The most common type of simultaneous transmission system is code-division multiplexing, where each user modulates a preassigned signature waveform known by the receiver.

Specifically, we will assume that in order to send the message $\{b_k(i) \in A\}_{i=0}^{M-1}$ (i.e., a string of M symbols drawn from a finite set A), the kth user transmits

$$\sum_{i=0}^{M-1} b_k(i) s_k(t - iT)$$

where $\{s_k(t), 0 \le t \le T\}$ is the waveform assigned to the kth user, and T is the symbol period. Then the demodulator receives the sum of the signals transmitted by the K active users embedded in noise

$$r(t) = \sum_{k=1}^{K} \sum_{i=0}^{M-1} b_k(i) s_k(t - iT - \tau_k) + n(t) \quad (3.8)$$

where the offsets $\tau_{k-1} \le \tau_k \in [0, T)$ model the fact that the users do not synchronize their transmissions. Then the task of the receiver is to recover the transmitted information strings $\{b_k(i)\}_{i=0}^{M-1}{}_{k=1}^{K}$. Following [47] we will show how to obtain an optimum multiuser demodulator via dynamic programming. First, denote the MK-vector

$$d = \{d_{k+iK} = b_k(i),\ k = 1, \cdots, K,\ i = 0, \cdots, M-1\}$$

and the multiuser signal in (3.8)

$$S(t, d) = \sum_{k=1}^{K} \sum_{i=0}^{M-1} b_k(i) s_k(t - iT - \tau_k) = \sum_{i=1}^{MK} d_i z_i(t) \quad (3.9)$$

where $z_{k+iK}(t) = s_k(t - iT - \tau_k)$.

A reasonable criterion for demodulating the information carried in $S(t, d)$ upon observation of $r(t)$ is to select the MK-vector d that best explains the received waveform in the sense of minimizing the energy of the corresponding noise realization, i.e.,

$$\min_{d \in A^{MK}} \| S(t, d) - r(t) \|^2. \quad (3.10)$$

If the noise $n(t)$ is white and Gaussian, then this criterion results in maximum likelihood decisions. Equivalently, the objective is to find the vector that solves

$$\max_{d \in A^{MK}} \Omega(d) \quad (3.11)$$

where

$$\Omega(d) = 2 \int_{-\infty}^{\infty} S(t, d) r(t)\, dt - \int_{-\infty}^{\infty} S^2(t, d)\, dt. \quad (3.12)$$

Since the maximization in (3.11) is over a finite set, we could solve it by the brute-force method of evaluating $\Omega(d)$ for each possible argument. However, it is possible to decompose $\Omega(d)$ in a sequential fashion that lends itself to efficient optimization. From

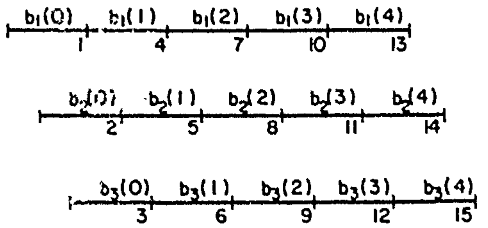


Fig. 6. Symbol epochs for $K = 3$ and $M = 5$.

(3.9) it is immediate to write the first integral in (3.12) sequentially

$$\int_{-\infty}^{\infty} S(t, d)r(t)\, dt = \sum_{j=1}^{MK} d_j y_j \tag{3.13}$$

where

$$y_j = \int_{-\infty}^{\infty} z_j(t) r(t)\, dt. \tag{3.14}$$

This implies that the objective function (3.12) depends on $r(t)$ only through the quantities $\{y_j\}_{j=1}^{MK}$, which are obtained by correlating $r(t)$ with each of the signature waveforms during each symbol epoch. In order to find an explicit expression for the second integral on the right-hand side of (3.12), which is the energy of the multiuser signal, we will denote

$$R(j, l) = \int_{-\infty}^{\infty} z_j(t) z_l(t)\, dt. \tag{3.15}$$

It follows immediately from the definition that these coefficients satisfy the following properties.

1) $R(k + iK, k + iK) = \int_0^T s_k^2(t) \triangleq w_k$.
2) $R(k + iK, n + iK) = R(k, n)$ for all i.
3) $R(j, l) = 0$ unless $|j - l| < K$.

The first property indicates that each of the diagonal elements of $R(i, j)$ is equal to the energy of one of the K assigned waveforms. The second and third properties can be illustrated by referring to Fig. 6 which represents the symbol epochs of three asynchronous users sending strings of $M = 5$ symbols. Each symbol period in Fig. 6 is labeled with the index of the corresponding component of the vector d. The second property indicates that the cross-correlations between two signals depend only on their relative location (e.g., $R(4, 6) = R(13, 15)$ in Fig. 6) and the third property states that each symbol only interferes with $2K - 2$ symbols of the other users [e.g., in Fig. 6, $d_9 = b_3(2)$ only overlaps with $d_7 = b_1(2)$, $d_8 = b_2(2)$, $d_{10} = b_1(3)$, and $d_{11} = b_2(3)$]. It follows from these properties that the coefficients in (3.15) can be obtained from the $K \times K$ matrix $\{R(k, n)\}_{k=1}^{K}{}_{n=1}^{K}$ whose diagonal elements correspond to the energy per symbol of each user and whose off-diagonal elements correspond to the cross-correlations between the signature waveforms of each pair of users. Using (3.15), the foregoing properties, and letting $\kappa(j) \in \{1, \cdots, K\}$ be the modulo-K remainder of j (i.e., for some i, $j = \kappa(j) + iK$), we can write

$$\begin{aligned}\int_{-\infty}^{\infty} S^2(t, d)\, dt &= \sum_{j=1}^{MK} \sum_{l=1}^{MK} d_j d_l R(j, l)\\ &= \sum_{j=1}^{MK} d_j \left[w_{\kappa(j)} + 2 \sum_{l=j-K+1}^{j-1} d_l R(j, l) \right]\\ &= \sum_{j=1}^{MK} d_j \left[w_{\kappa(j)} + 2 \sum_{n=1}^{K-1} d_{j-n} g_{\kappa(j)}(K-n) \right]\end{aligned} \tag{3.16}$$

where $g_k(m) = R(k + K, k + m)$. Putting together (3.12), (3.13), and (3.16) we see that we can express $\Omega(d)$ as a sum of MK terms, each of which depends on K components of d and such that consecutive terms depend on the same components but one. Specifically, we can write

$$\Omega(d) = \sum_{j=1}^{MK} \lambda_j(x_j, d_j) \tag{3.17}$$

where

$$\lambda_j(x, u) = u[2y_j + u w_{\kappa(j)} - 2x^T g_{\kappa(j)}] \tag{3.18}$$

and x_j is the state of a shift-register $K - 1$ dimensional system

$$x_{j+1}^T = [x_{j+1}(1), \cdots, x_{j+1}(K-1)] = [x_j(2), \cdots, x_j(K-1), d_j];$$
$$x_0 = 0. \tag{3.19}$$

It is now apparent that the solution to (3.11) entails solving a *finite-horizon deterministic optimal control problem* with additive costs per stage for the linear system in (3.19), and with a finite admissible control set A. Therefore, optimum multiuser demodulation is equivalent to a shortest path problem in an M-stage layered directed graph, where at each stage there are A^{K-1} states. This optimization problem can be solved by dynamic programming (e.g., [7]) in backward or forward fashion. In practice, it is necessary to demodulate the transmitted symbols in real-time, and since M is usually a very large integer, it is not feasible to wait until all the observables $\{y_j\}_{j+1}^{MK}$ have been obtained before starting to make decisions. Therefore, a suboptimum version of the forward dynamic programming algorithm is adopted in practice whereby each decision is based on the paths corresponding to the cost-to-arrive function computed a fixed number of steps ahead. This real-time version of forward dynamic programming is known in communication theory as the Viterbi algorithm [12], and was originally devised (without resorting to the dynamic programming framework) for decoding convolutional codes. The maximum-likelihood criterion used in (3.10) is not the only possible optimality criterion. For example, if the objective is to minimize the probability of error for each user, then the multiuser demodulator uses a *backward-forward* dynamic programming algorithm [49] whereby optimum decisions are based on the independent computation of a cost-to-go and a cost-to-arrive function.

IV. Other Problem Areas

Routing and multiple access are not the only problem areas in the field of communication networks which control theory can help formulate, study, and solve. We have deliberately chosen to confine our attention to these two areas in order to get across in a concise manner our belief that the field of communication networks offers a rich selection of applications for control theory. We would feel remiss, however, if we did not even make an attempt to provide a taste of some of the numerous other design and operation issues that, again, bring forth control systems concepts and techniques. For this purpose, and with a conscious effect not to expand in depth but only to describe, we will mention two areas from point-to-point networks and one from radio networks. The first two concern flow control and integrated switching, respectively, while the third concerns the problem of scheduling transmission in multihop networks. Unlike the cases of routing and multiple access, these areas have not yet fully benefitted from the use of control theoretic approaches although such approaches would be very well suited to them indeed.

A. Flow Control

A stark reality in the design of networks is that despite the reduction of the cost of memory, storage at each node is going to be finite. Coupled with another reality, namely that data transmissions on the whole continue to be bursty, it implies that buffer overflow may occur and, along with it, congestion and deadlocks. Flow control is the name we use to describe the collection of measures taken to avoid buffer overflow and highly congested nodes in the network. Congestion and saturation are often the consequences of diverging, unstable behavior. Thus, it is of interest not only to optimize over possible flow control strategies, but to determine their robustness against disturbances or modeling inaccuracies that may lead to unstable behavior.

The control variables in flow control problems are admission (or blocking) probabilities for messages or sessions at the source node. In practice these are often implemented in terms of a bang-bang control strategy known as *window flow control* whereby input ports are allowed to continuously inject messages into the network at the full desired input rate until the number of *unacknowledged*[6] messages exceeds the value of the "window size" w. A simple, yet unanswered question is, what should the value of w be?

Previous efforts to use control theory tools to analyze optimal flow control problems include [28] and [46] where the optimality of window flow control is proved within the domain of a simplified model, and [39] where dynamic programming value iteration techniques are used to characterize optimal flow control performance. An alternative approach to the flow control problem is to subsume it into the static routing problem considered in Section II-A [19]: suppose that for every source-destination pair a fictitious direct link is added between them. We can then interpret the blocking action of a flow control procedure as a diversion of the blocked portion of the traffic through this fictitious link to the destination. Thus, we can consider that no traffic is blocked. Of course, in order to discourage the use of this fictitious link we must augment the overall delay cost function with a term that penalizes appropriately the use of this link.

B. Integrated Switching

A revolutionary development in the field of networks whose implementation is currently under way is the combination of the capabilities of what have been separately developed in the past and called voice networks and data networks. Voice is a commodity that must meet different requirements than data. For example, speech signals have inherent redundancy that make them quite robust with respect to occasional errors or deliberate compression. At the same time, except in applications of voice messaging, speech signals occur in the context of real-time conversations and, as such, must encounter short and, more importantly, constant delay. On the other hand, data must preserve their integrity and cannot tolerate errors, however, long and variable delays can be often tolerated.

How does one design a single network that can handle such dissimilar commodities with automated procedures? The natural course of events in the last decade or two was to attempt to force data on primarily voice networks or to let voice ride on what were mainly data networks. The literature is full of ideas for baseline integration that are mostly heuristic and difficult to analyze. An attempt to formulate the problem of integrated switching as an optimization problem was presented in [50]. In it simplest form the model is as follows: consider a single node in the network with a single outgoing link on which incoming voice calls and data packets must be multiplexed. Let W be the bandwidth of the outgoing link. Let V be the bandwidth required for the continuous, uninterrupted accommodation of a single voice call. Let, therefore, $N = W/V$ be the maximum number of calls that can be assigned dedicated circuits simultaneously if no data packets are transmitted. A voice call can either be accepted (and assigned the necessary bandwidth V) or blocked. Data packets can be stored in a buffer facility. If, at a given time, there are i calls in the system, the data packets can be served at the full rate corresponding to the remaining bandwidth $W - iV$. Such a switching architecture represents what has been called the *movable boundary* idea in integration. A natural MDP can be simply formulated as follows: choose the control action of blocking or accepting a call upon arrival in order to minimize the weighted sum of the average data packet delay and the call-blocking probability. If we assume that both arrival streams (voice calls and data) are independent Poisson processes, that the call holding time is exponentially distributed, and that the message lengths are likewise exponential, we can apply the technique described in Section II of converting the MDP to an LP and show that the optimal policy has the useful *switching-type* form. Namely, if i is the number of ongoing calls and j the total number of data messages at the node, the optimal control action should be to block the call in region B of the state space as shown in Fig. 7 and to accept it in region A.

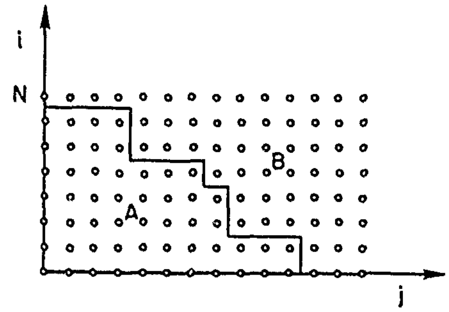


Fig. 7. Switching-type optimum policy for integrated switching.

C. Link Scheduling

Let us now turn our attention back to the radio network environment. In Section III the multiple access channel was considered and a number of difficult but interesting control problems were identified. Throughout that discussion, it was assumed that all terminals are within a single transmission hop from the destination. In many radio networks, however, this is not the case. Messages need to be relayed via intermediate nodes to their final destinations. Thus, the familiar problem of routing arises again, except that this time there is a new twist to it. In point-to-point networks, transmissions between different node pairs can take place simultaneously because there are dedicated, "hard-wired" links between the corresponding nodes. In a radio (or, more generally, in a multiaccess/broadcast) environment, if the nodes are densely connected, not all transmissions can take place simultaneously (unless separate dedicated channels or simultaneous transmission signaling techniques (Section III-B) are used). They must be scheduled in time to avoid the interference that would occur otherwise.

It becomes evident that the mere fact that the transmission among a group of nodes must take place one at a time raises the question whether the intended transmissions are routing-wise optimal any more. Several versions of this problem have been studied [3], [23], [36]. In every case and even if the routing problem is sidestepped, we are led to hard combinatorial optimization problems where questions of computational complexity and distributed implementation are of primary importance.

V. Conclusion

It should be clear by now that the theory of linear and nonlinear optimization, dynamic programming, stochastic control, stability

[6] Note the implicit assumption of delayed feedback information from the destination to the source node.

analysis, and distributed control have found interesting applications arising in the analysis and design of communication networks. Unlike other complex systems that have been successfully studied by control system theorists in the past (such as chemical plants, flexible aircraft, robot systems, etc.), communication networks stand out in that the commodity to be controlled is information (including its transmission, storage, processing, etc.). This feature, perhaps, misleads and intimidates those who do not feel sufficiently inter-disciplinarian to tackle these problems. We hope that by having selected to present a few examples in which concrete, purely control-theoretic problems can be formulated and have been (or can be) studied successfully, we may encourage attention by the control community to this application area that is especially rich in new challenges.

As stated from the outset, we did not attempt to survey or completely cover the multiple control facets of communication networks. The collection in this paper merely represents an effort to illuminate a few selected problem areas and to show how control techniques apply to them.

References

[1] N. Abramson, "Development of the ALOHANET," *IEEE Trans. Inform. Theory*, vol. IT-31, pp. 119-123, Mar. 1985.

[2] P. O. Bhattacharya *et al.* "A (not very) simple dynamic routing problem," in *Proc. 25th Allerton Conf. Contr. Commun. Comput.*, Urbana, IL, 1987, pp. 998-1006.

[3] D. J. Baker, J. Wieselthier, and A. Ephremides, "A distribured algorithm for scheduling the activation of links in a self-organizing mobile radio network," in *Proc. IEEE Int. Conf. Commun.*, Philadelphia, PA, June 1982, pp. 2F.6.1-5.

[4] M. Bazaraa and C. Shetty, *Nonlinear Programming: Theory and Algorithms*. New York: Wiley, 1979.

[5] D. Bertsekas and R. Gallager, *Data Networks*. Englewood Cliffs, NJ: Prentice-Hall, 1987.

[6] D. P. Bertsekas, "Distributed dynamic programming," *IEEE Trans. Automat. Contr.*, vol. AC-27, pp. 610-616, June 1982.

[7] D. P. Bertsekas, *Dynamic Programming: Deterministic and Stochastic Models*. Englewood Cliffs, NJ: Prentice-Hall, 1987.

[8] C. Buyukkoc, P. Varaiya, and J. Walrand, "The μc rule revisited," *Advances Appl. Probability*, vol. 17, pp. 237-238, 1985.

[9] C. A. Courcoubetis, "Optimal control of a queueing system with simultaneous service requirements," *IEEE Trans. Automat. Contr.*, vol. AC-32, pp. 717-727, 1987.

[10] A. Ephremides, P. Varaiya, and J. Walrand, "A simple dynamic routing problem," *IEEE Trans. Automat. Contr.*, vol. AC-25, pp. 690-693, Aug. 1980.

[11] G. Fayolle, E. Gelenbe, and L. Labetoulle, "Stability and optimal control of the packet switching broadcast channel," *JACM*, vol. 24, pp. 375-386, 1977.

[12] G. D. Forney, "The Viterbi algorithm," *Proc. IEEE*, vol. 61, pp. 268-278, Mar. 1973.

[13] M. Frank and P. Wolfe, "An algorithm for quadratic programming," *Naval Res. Logist. Quat.*, vol. 3, pp. 149-154, 1956.

[14] L. Fratta, M. Gerla, and L. Kleinrock, "The flow deviation method: An approach to store-and-forward communication network design," *Networks*, vol. 3, pp. 97-133, 1973.

[15] T. Gal, *Postoptimal Analyses, Parametric Programming, and Related Topics*. Englewood Cliffs, NJ: McGraw-Hill, 1979.

[16] R. Gallager, "A minimum delay routing algorithm using distributed computation," *IEEE Trans. Commun.*, vol. COM-23, pp. 73-85, 1977.

[17] S. Ghez, S. Verdu, and S. Schwartz, "Stability properties of slotted Aloha with multipacket reception capability," *IEEE Trans. Automat. Contr.*, vol. 33, pp. 640-649, July 1988.

[18] S. Ghez, S. Verdu, and S. Schwartz, "Optimal decentralized control in the multipacket channel," *IEEE Trans. Automat. Contr.*, to be published.

[19] S. J. Golestaani, "A unified theory of flow control and routing on data communication networks," Ph.D. dissertation, Mass. Inst. Technol., Cambridge, 1980.

[20] B. Hajek and T. van Loon, "Decentralized dynamic control of a multiaccess broadcast channel," *IEEE Trans. Automat. Contr.*, vol. AC-27, pp. 559-569, June 1982.

[21] B. Hajek, "Hitting-time and occupation-time bounds implied by drift analysis with applications," *Adv. Appl. Prob.*, vol. 14, pp. 502-525, 1982.

[22] ——, "Optimal control of two interacting service stations," *IEEE Trans. Automat. Contr.*, vol. AC-29, pp. 491-499, June 1984.

[23] B. Hajek and G. Sasaki, "Link scheduling in polynomial time," *IEEE Trans. Inform. Theory*, vol 34, pp. 910-917, Sept. 1988.

[24] J. M. Harrison, "Dynamic scheduling of a multiclass queue. Discount optimality," *Oper. Res.*, vol. 23, pp. 270-282, 1975.

[25] F P Kelly, "Stochastic models of computer communications systems," *J. R. Statist. Soc. B*, vol. 47, pp. 379-395, 1985.

[26] G. P. Klimov, "Time sharing service systems," *Theory Probability Appl.*, vol. 19, pp. 532-551, Sept. 1974, see also vol. 23, pp. 314-321, June 1978.

[27] H. Kushner, *Introduction to Stochastic Control*. New York, Holt, Rinehart, and Winston, 1971.

[28] A. Lazar, "Optimum flow control of a class of queueing networks in equilibrium," *IEEE Trans. Automat. Contr.*, vol. AC-28, pp. 1001-1007, Nov. 1983.

[29] W. Lin and P. R. Kumar, "Optimal control of a queueing system with two heterogeneous servers," *IEEE Trans. Automat. Contr.*, vol. AC-29, pp. 696-703, Aug. 1984.

[30] S. A. Lippman, "Semi-Markov decision processes with unbounded rewards," *Management Sci.*, vol. 19, pp. 717-731, 1973.

[31] V. A. Mikhailov, "Geometrical analysis of the stability of Markov chains in R^n_+ and its application to throughput evaluation of the adaptive random multiple access algorithm," *Problems of Inform. Transmission*, vol. 24, pp. 47-56, Jan.-Mar. 1988.

[32] J Mosely and P Humblet, "A class of efficient contention resolution algorithms for multiple access channels," *IEEE Trans. Commun.*, vol. COM-33, pp.145-151, Feb. 1985.

[33] P. Nain and K. W. Ross, "Optimal priority assignment with hard constraint," *IEEE Trans. Automat. Contr.*, vol. AC-31, pp. 883-888, 1986.

[34] C. H. Papadimitrou and K. Steiglitz, *Combinatorial Optimization: Algorithms and Complexity*. Englewood Cliffs, NJ. Prentice-Hall, 1982.

[35] S. Parekh, F. Schoute, and J. Walrand, "Instability and geometric transience of the Aloha protocol," in *Proc. 26th Conf. Decision Contr.*, Los Angeles, CA, Dec. 1987, pp. 1073-1077.

[36] M. J. Post, P. E. Sarachik, and A. S. Kershenbaum, "A biased greedy algorithm for scheduling multi-hop radio networks," in *Proc. Conf. Inform. Sci. Syst.*, Johns Hopkins Univ., Baltimore, MD, 1985, pp. 564-672.

[37] R. T. Rockafellar, *Convex Analysis*. Princeton, NJ. Princeton University Press, 1970.

[38] Z. Rosberg, P. Varaiya, and J. Walrand, "Optimal control of service in tandem queues," *IEEE Trans. Automat. Contr.*, vol. AC-27, pp. 600-610, June 1982.

[39] Z. Rosberg and I. Gopal, "Optimal hop-by-hop flow control in computer networks," *IEEE Trans. Automat. Contr.*, vol. AC-31, pp. 813-822, Sept. 1986.

[40] W. Rosberg and D. Towsley, "On the instability of slotted-ALOHA multiaccess algorithm," *IEEE Trans. Automat. Contr.*, vol. AC-28, pp. 994-996, Oct. 1983.

[41] R. E. Tarjan, *Data Structures and Network Algorithms* (CBMS-NSF Reg. Conf. Series in Appl. Math. no. 7). Philadelphia, PA: SIAM, 1983.

[42] J. N. Tsitsiklis and D. P. Bertsekas, "Distributed asynchronous optimal routing in data networks," *IEEE Trans. Automat. Contr.*, vol. AC-31, pp. 325-332, Apr. 1986.

[43] J. N. Tsitsiklis, "Analysis of a multiacess control scheme," *IEEE Trans. Automat. Contr.*, vol. AC-32, pp. 1017-1020, Nov. 1987.

[44] B. S. Tsybakov and N. B. Likhanov, "An improved upper bound on capacity of the random multiple-access channel," *Problemi Pederachi Informatsii*, vol. 23, pp. 64-78, 1987.

[45] R. Tweedie, "Sufficient conditions for ergodicity and recurrence of Markov chains on a general state space," *Stoch. Proc. Appl.*, vol. 3, pp. 385-403, 1975.

[46] F. Vakil and A. Lazar, "Flow control protocols for integratred networks with partially observed voice traffic," *IEEE Trans. Automat. Contr.*, vol. AC-32, pp. 2-14, Jan. 1987.

[47] S. Verdu, "Minimum probability of error for asynchronous Gaussian multiple-access channels," *IEEE Trans. Inform. Theory*, vol. IT-32, pp. 85-96, Jan. 1986.

[48] —, "Computation of the efficiency of the Mosley-Humblet contention resolution algorithm. A simple method," *Proc. IEEE*, vol. 74, pp. 613-614, Apr. 1986.

[49] S. Verdu and H. V. Poor, "Abstract dynamic programming models under commutativity conditions," *SIAM J. Contr. Optimiz.*, vol. 4, July 1987.

[50] I. Viniotis and A. Ephremides, "Optimal switching of voice and data at a network node," *Proc. 26th CDC*, Los Angeles, CA, Dec. 1987, pp. 1504-1507.

[51] J. Viniotis, Ph.D. dissertation, Univ. Maryland, College Park, MD, 1988.

[52] J. Walrand, "A note on optimal control of a queueing system with two heterogeneous servers," *Syst. Contr. Lett.*, vol. 4, pp. 131-134, 1984.

[53] ——, *An Introduction to Queueing Networks*. Englewood Cliffs, NJ: Prentice-Hall, 1988.

[54] R. Weber, "On the optimal assignment of customers to parallel servers," *J. Appl. Prob.*, vol. 15, pp. 406-413, 1978.

[55] Z. Wu, P. B. Luh, S. Chang, and D. A. Castanon, "Optimal control of

a queueing system with two interacting service stations and three classes of impatient tasks," *IEEE Trans. Automat. Contr.*, vol. 33, pp. 42–49, 1988.

Anthony Ephremides (S'68–M'71–SM'77–F'84) was born in Athens, Greece, in 1943. He received the Ph.D. degree in electrical engineering from Princeton University, Princeton, NJ, in 1971.

He has been with the University of Maryland, College Park, since 1971. He has also spent semesters on leave at M.I.T., the University of California, Berkeley, and ETH, Zurich. He is active in professional consulting as President of Pontos, Inc. Currently, his research interests lie in the areas of communication systems: performance analysis; modeling; optimization; simulation; and design.

Dr. Ephremides is Director of Division X of the IEEE and has served as President of the Information Theory Society and on the Board of Governors of the Control Systems Society. He has been Associate Editor of the IEEE TRANSACTIONS ON AUTOMATIC CONTROL and General Chairman of major IEEE Conferences.

Sergio Verdú (S'80–M'84–SM'88) was born in Barcelona, Catalonia, Spain, in 1958. He received the Telecommunication Eng. degree from the Polytechnic University of Barcelona in 1980 and the Ph.D. degree in electrical engineering from the University of Illinois at Urbana-Champaign in 1984.

Upon completion of his doctorate he joined the faculty of Princeton University, Princeton, NJ, where he is currently an Associate Professor of Electrical Engineering. His current research interests are in the areas of multiuser communication and information theory.

Dr. Verdú is a recipient of the National University Prize of Spain, the Rheinstein Outstanding Junior Faculty Award of the School of Engineering and Applied Science at Princeton University, and the NSF Presidential Young Investigator Award. He is currently serving as Associate Editor of the IEEE TRANSACTIONS ON AUTOMATIC CONTROL, and as a member of the Board of Governors of the IEEE Information Theory Society.

T-AC/34/11//30788

Optimal Decentralized Control in the Random Access Multipacket Channel

Sylvie Ghez
Sergio Verdú
Stuart C. Schwartz

Reprinted from
IEEE TRANSACTIONS ON AUTOMATIC CONTROL
Vol. 34, No. 11, November 1989

Optimal Decentralized Control in the Random Access Multipacket Channel

SYLVIE GHEZ, STUDENT MEMBER, IEEE, SERGIO VERDÚ, SENIOR MEMBER, IEEE, AND STUART C. SCHWARTZ, SENIOR MEMBER, IEEE

Abstract—**A decentralized control algorithm is sought that maximizes the stability region of the infinite-user slotted multipacket channel and is easily implementable. To this end, the perfect state information case where the stations can use the instantaneous value of the backlog to compute the retransmission probability is studied first. The best throughput possible for a decentralized control protocol is obtained, as well as an algorithm that achieves it. Those results are then applied to derive a control scheme when the backlog is unknown, which is the case of practical relevance. This scheme, based on a binary feedback, is shown to be optimal given some restrictions on the channel multipacket reception capability.**

I. Introduction

Most studies on random access communications rely on the assumption that when two or more packets overlap, all the information that was sent is irremediably lost, hence the need to repeat all transmissions at some later time. This is actually a pessimistic point of view, since there are many examples of random access systems where one or more packets may be successful in the presence of other simultaneous transmissions. In order to represent such random access systems, a model for a channel with multipacket reception capability has been developed in [6]–[8]. We consider a slotted channel with an infinite population of users, and we assume that the probability of having k successes in a slot where there are n transmissions depends only on the collision size n

$$\epsilon_{nk} = P[k \text{ packets are correctly received} \mid n \text{ are transmitted}] \qquad (n \geq 1,\ 0 \leq k \leq n).$$

We define the reception matrix as

$$E = \begin{bmatrix} \epsilon_{10} & \epsilon_{11} & & & \\ \epsilon_{20} & \epsilon_{21} & \epsilon_{22} & 0 & \\ \cdot & \cdot & & & \\ \epsilon_{n0} & \epsilon_{n1} & & \epsilon_{nn} & \\ \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & & & \cdot \end{bmatrix}.$$

This model can be applied to channels with capture [1]–[3], [10], [16], [18], [20], [23], [26], [28], [34] and to systems using CDMA [22], [24], [29]. It is also relevant for many other applications, such as systems with multiuser detectors [33] or, for instance, the channel studied in [17], [31]. For more details about this model, the reader is referred to [6] and [8]. Denoting by $C_n = \sum_{k=1}^{n} k\epsilon_{nk}$ the average number of packets correctly received in collisions of size n, we assume that the limit $C = \lim_{n\to\infty} C_n$ exists, as is usually the case with models of practical interest. It has been proved in [8] that the Aloha random access algorithm has a maximum stable throughput $\eta_0 = C$ in the multipacket channel.

Decentralized control strategies have been shown [11], [12], [19], [25], [30] to stabilize the slotted Aloha algorithm in the case of the usual collision channel, hence, it is reasonable to expect that when those strategies are used in the multipacket channel, the resulting throughput will be higher than η_0. We consider schemes of the form

$$p_n = F(S_n)$$
$$S_{n+1} = G(S_n, Z_n) \tag{1}$$

where p_n is the retransmission probability in slot n, S_n is an estimate of the backlog X_n at the beginning of slot n, and Z_n is the feedback at the end of slot n. The number of new packets arriving during slot n, A_n is assumd to form a sequence of i.i.d. random variables with probability distribution $P[A_n = k] = \lambda_k (k \geq 0)$, such that the mean arrival rate $\lambda = \sum_{n=1}^{\infty} n\lambda_n$ is finite. Each of the A_{n-1} new packets that arrived during slot $n - 1$ is transmitted in slot n with probability p_n.

As in the case of conventional channels, it is useful to study first the case of control with perfect state information where the value of the backlog is given to the users prior to the selection of the retransmission probability. To keep track of the exact value of the backlog, a central controller is usually necessary, which is an unreasonable requirement for most practical random access channels. However, the study of the perfect state information case allows us to determine an upper bound to the best throughput η_c achievable by any decentralized control of the form (1), and suggests a simple implementation. Those results are in turn helpful to derive control protocols in the case where the backlog is unknown. This is done in Section III where we consider a backlog estimate which is recursively updated using the binary feedback empty/nonempty. In addition, it is assumed throughout the paper that each station is informed when its packet is successfully received. It is proved that provided a certain condition on the reception matrix holds, the throughput achievable with this type of feedback is the same as the perfect state information throughput. This condition is verified for most multipoint-to-point channels of practical interest.

In a paper whose translation appeared only very recently [19] (after our work [7]), Mikhailov has derived sufficient conditions for stability and instability of two-dimensional Markov chains. Although this was meant to be used for decentralized control schemes in the usual collision channel, this approach is powerful enough to be applied to the multipacket channel. In Section IV we show by using Mikhailov's result that the scheme presented in Section III is stable under weaker assumptions. However, only a weaker form of stability can be proved in this way.

Manuscript received April 18, 1988; revised January 20, 1989 and May 7, 1989. Paper recommended by Associate Editor, X. R. Cao. This work was supported in part by the Office of Naval Research under Contracts N00014-87-K-0054, by the Army Research Office under Contract DAAL-03-87-K-0062, and by the New Jersey Commission on Science and Technology under Grant 85-990660-6

The authors are with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544.

IEEE Log Number 8930788.

0018-9286/89/1100-1153$01.00

II. Control of the Multipacket Channel with Perfect State Information

In this section we assume that all the users know the value of X_n at the beginning of slot n, and we let the retransmission probability be a function of the exact value of the backlog, i.e., $p_n = F(X_n)$. In this ideal case, the system is much simpler to analyze than in the general case (1) since $(X_n)_{n\geq 0}$ is a homogeneous Markov chain. Our goal is to determine the optimal control function F^* that yields the largest ergodicity region, and the corresponding throughput, denoted by η_c. For instance, it is well known [4] that for the usual collision channel with the access rule in effect here, $F^*(X_n) = 1/X_n$ is the retransmission probability that minimizes the drift at each step, resulting in an ideal throughput of $\eta_c = e^{-1}$.

First note that all the results herein are valid provided that the backlog Markov chain $(X_n, S_n)_{n\geq 0}$ corresponding to a control (1) is irreducible and aperiodic. It can be easily checked that for both access rules considered in this paper (see below), as well as all the algorithms, a simple set of sufficient conditions for irreducibility and aperiodicity is

a) $\lambda_0 \neq 0$

b) $\lambda_0 + \sum_{n=1}^{\infty} \lambda_n \epsilon_{nn} < 1$

c) $\epsilon_{10} \neq 0$

which are analogous to the conditions for the open-loop system studied in [6]. The theorem below gives the best throughput possible for a control protocol (1).

Theorem 1: There exists a retransmission probability p_n^* that minimizes the expected backlog increase when the backlog is equal to n.

With such a retransmission probability, the system is stable for $\lambda < \eta_c$ and unstable for $\lambda > \eta_c$, with

$$\eta_c = \sup_{x\geq 0} e^{-x} \sum_{n=1}^{\infty} C_n \frac{x^n}{n!}.$$

Proof of Theorem 1: The proof is based on standard drift analysis techniques. $(X_t)_{t\geq 0}$ is a homogeneous Markov chain which evolves according to

$$X_{t+1} = X_t + A_t - \Sigma_t \tag{2}$$

where Σ_t is the number of packets successfully transmitted in slot t. The system is defined to be stable if $(X_t)_{t\geq 0}$ is ergodic and unstable otherwise. Let d_n be the drift of X_t at state n: $d_n = E[X_{t+1} - X_t | X_t = n]$. We have $0 \leq \Sigma_t \leq X_t$, and if we denote by p the retransmission probability used in slot t, then for $n \geq 1$, the probability of having k successes is given by

$$P[\Sigma_t = k | X_t = n] = \sum_{j=k}^{n} \binom{n}{j} p^j (1-p)^{n-j} \epsilon_{ij} \qquad (1 \leq k \leq n). \tag{3}$$

It then follows from (2) that the backlog drift at state $n \geq 1$ is given by

$$\begin{aligned} d_n &= \lambda - \sum_{k=1}^{n} k \sum_{j=k}^{n} \binom{n}{j} p^j (1-p)^{n-j} \epsilon_{jk} \\ &= \lambda - \sum_{j=1}^{n} \binom{n}{j} p^j (1-p)^{n-j} C_j \end{aligned} \tag{4}$$

which becomes $d_n(p) = \lambda - t_n(p)$ if we define $t_n(p)$ to be the average number of successes given the backlog n and the retransmission probability p

$$t_n(p) = \sum_{j=1}^{n} \binom{n}{j} p^j (1-p)^{n-j} C_j. \tag{5}$$

Since $t_n(p)$ is a polynomial on the compact $[0,1]$, it achieves its maximum and we can define

$$p_n^* = \arg\max_{p\in[0,1]} t_n(p) = \arg\min_{p\in[0,1]} d_n(p).$$

We now proceed to compute the limit of the drift when the retransmission probability p_n^* is used. We show that

$$\lim_{n\to\infty} t_n(p_n^*) = \sup_{x\geq 0} e^{-x} \sum_{n=1}^{\infty} C_n \frac{x^n}{n!} = \sup_{x\geq 0} t(x). \tag{6}$$

Let us first assume that $C < +\infty$.

Property 1:

$$\lim_{x\to\infty} t(x) = C.$$

We have for $n > M$

$$|t(x) - C| \leq e^{-x} C + e^{-x} \sum_{n=1}^{M} \frac{x^n}{n!} |C_n - C| + \sum_{n=M+1}^{\infty} \frac{x^n}{n!} |C_n - C|. \tag{7}$$

Pick $\epsilon > 0$ and fix M such that $|C_n - C| < \epsilon$ for $n > M$. Then if B_c is an upper bound on the sequence $(C_n)_{n\geq 1}$, (7) yields

$$|t(x) - C| \leq e^{-x} C + 2B_c e^{-x} \sum_{n=1}^{M} \frac{x^n}{n!} + \epsilon$$

and the right-hand side of this last equation goes to zero as x goes to infinity.

Property 2: For all $\epsilon > 0$, there exists $A > 0$ such that for all $np > A$, $|t_n(p) - C| < \epsilon$. We have

$$|t_n(p) - C| \leq \sum_{j=1}^{n} \binom{n}{j} p^j (1-p)^{n-j} |C_j - C| + (1-p)^n C.$$

Choosing M as for Property 1 we get

$$|t_n(p) - C| \leq 2B_c \sum_{j=0}^{n} \binom{n}{j} p^j (1-p)^{n-j} + \epsilon.$$

Let us denote by R_n the random variable corresponding to the number of retransmissions in a slot given that the backlog is equal to n. We have

$$\sum_{j=0}^{M} \binom{n}{j} p^j (1-p)^{n-j} = P[R_n \leq M] \leq P\left[\left|\frac{R_n}{n} - p\right| > \frac{p}{2}\right]$$

for $np > 2M$. Then from the Chebyshev inequality

$$P[R_n \leq M] \leq \frac{4}{np} \tag{8}$$

and Property 2 follows.

Property 3: $t_n(x/n)$ converges uniformly to $t(x)$ on any compact $[0, A]$.

Fix $\epsilon > 0$ and choose M such that $\Sigma_{j=M}^{\infty} A^j C_j/j! < \epsilon$. Then for $n > M + 1$ and $x \in [0, A]$

$$\left|t_n\left(\frac{x}{n}\right)-t(x)\right| \le \sum_{j=1}^{M} A^j \frac{C_j}{j!}\left| e^{-x} - \frac{n(n-1)\cdots(n-j+1)}{n^j}\left(1-\frac{x}{n}\right)^{n-j}\right| + 2\epsilon.$$

Since $\lim_{n\to\infty} n(n-1)\cdots(n-j+1)/n^j = 1$ for $1 \le j \le M$, it is enough to show that $(1 - x/n)^{n-j}$ converges uniformly to e^{-x} for $1 \le j \le M$. We have

$$\left(1-\frac{x}{n}\right)^{n-j} - e^{-x} \le e^{-x}[e^{xj/n}-1] \le e^{AM/n}-1. \tag{9}$$

On the other hand, for $n > A$,

$$\left(1-\frac{x}{n}\right)^{n-j} - e^{-x} \ge \left(1-\frac{x}{n}\right)^{n} - e^{-x} \ge e^{-x}[e^{A+n\log(1-A/n)}-1]$$
$$\ge e^{A}\left(1-\frac{A}{n}\right)^{n}-1 \tag{10}$$

and uniform convergence follows from (9) and (10).

Property 4: $t_n(x/n)$ converges uniformly to $t(x)$ for $x \ge 0$.

Fix $\epsilon > 0$. From Properties 1 and 2 we can fix A such that:
i) for all $np > A$, $|t_n(p) - C| < \epsilon$,
ii) for all $x > A$, $|t(x) - C| < \epsilon$.

Then we distinguish two cases. If $x \in [0, A]$, then from Property 3 there exists N such that for all $n \ge N$, $|t_n(x/n) - t(x)| < \epsilon$. If on the other hand $x \in (A, +\infty)$, we have

$$\left|t_n\left(\frac{x}{n}\right)-t(x)\right| \le \left|t_n\left(\frac{x}{n}\right)-C\right| + |t(x)-C| \le 2\epsilon \tag{11}$$

from i) and ii).

Thus, we have shown that when C is finite, $t_n(x/n)$ converges uniformly to $t(x)$ for $x \ge 0$. It follows that $\lim_{n\to\infty}\sup_{x\ge 0} t_n(x/n) = \sup_{x\ge0} t(x)$ and so (6) is proved.

Finally, we show that (6) holds when $C = +\infty$. Choose Δ arbitrarily large and M such that $C_n > \Delta$ for $n > M$. Then for $n > M$

$$t_n\left(\frac{x}{n}\right) \ge \Delta \sum_{j=M+1}^{n}\binom{n}{j}\left(\frac{x}{n}\right)^j\left(1-\frac{x}{n}\right)^{n-j} \ge \Delta(1-P[R_n \le M]).$$

From (8) $P[R_n \le M]$ is arbitrarily small for $nx/n = x$ large enough. Therefore, $\sup_{x\ge0} t_n(x/n) = +\infty$ and $\lim_{n\to\infty} t_n(p_n^*) = +\infty$. Since it is clear that if $C = +\infty$, then $\sup_{x\ge0} t(x) = +\infty$, (6) holds.

From the equality $\lim_{n\to\infty} d_n(p_n^*) = \lambda - \sup_{x\ge0} t(x)$ and Pakes Lemma in [21], it follows that if $\lim_{n\to\infty} C_n = +\infty$, then $\lim_{n\to\infty} d_n(p_n^*) = -\infty$, and the system is always stable, whereas if $\lim_{n\to\infty} C_n < +\infty$, then $(X_n)_{n\ge0}$ is ergodic for $\lambda < \eta_c = \sup_{x\ge0} t(x)$. Also, it is shown in the Appendix that Kaplan's condition holds for this system when the sequence $(C_n)_{n\ge1}$ is bounded, thus from Kaplan's result [13], the backlog Markov chain is nonergodic when $\lambda > \eta_c$. □

It is intuitively obvious that no decentralized control algorithm of the form (1) can have a maximum stable throughput larger than η_c. The theorem below gives a rigorous proof of this fact and also shows that this throughput can be achieved with a control which is much simpler than p_n^*.

Theorem 2: The best throughput achievable by a decentralized control algorithm (1) is $\eta_c = \sup_{x\ge0} e^{-x}\Sigma_{n=1}^{\infty} x^n/n!\, C_n$. If $\eta_c > C = \lim_{n\to\infty} C_n$, then there exists a constant $A > 0$ such that the control $p_t = A/X_t$ for $X_t > A$ yields the optimal throughput η_c.

Proof of Theorem 2: To prove the first part of the theorem we use a result of [27] which is a generalization of Kaplan's Theorem. If $p_t = F(S_t)$ and $S_{t+1} = G(S_t, Z_t)$, consider the Markov chain (X_t, S_t) and the Lyapunov function $V(n, s) = n$. Assume that $\lambda > \eta_c$. Then

$$E[V(X_{t+1}, S_{t+1}) - V(X_t, S_t) | X_t = n, S_t = s]$$
$$= \lambda - \sum_{j=1}^{n}\binom{n}{j} F(s)^j(1-F(s))^{n-j}C_j$$
$$\ge d_n(p_n^*) \ge \frac{\lambda-\eta_c}{2} \tag{12}$$

for all n large enough and all s. Therefore, the drift of V is strictly positive outside a finite subset of the state space. Since it is shown in the Appendix that the generalized Kaplan's condition is verified, it is enough to conclude that (X_t, S_t) is nonergodic. Hence, η_c is indeed the best throughput achievable by any decentralized control algorithm of the form (1).

To prove the second part of the theorem, we need the following property.

Property 5: If for all $x \ge 0$, $t(x) < \sup_{x\ge0} t(x)$, then $\sup_{x\ge0} t(x) = C$.

If $\sup_{x\ge0} t(x) = +\infty$, it is easily seen that $C = +\infty$. If $\sup_{x\ge0} t(x) < +\infty$, then $C < +\infty$. Consider a sequence $(x_n)_{n\ge1}$ of nonnegative reals such that $\lim_{n\to\infty} t(x_n) = \sup_{x\ge0} t(x)$. If $(x_n)_{n\ge1}$ was bounded above by $K < +\infty$, we would have for all $n \ge 1$, $t(x_n) \le \sup_{x\in[0,K]} t(x)$, and in the limit $\sup_{x\ge0} t(x) = \sup_{x\in[0,K]} t(x)$. Then there would exist $x_0 \in [0, K]$ such that $t(x_0) = \sup_{x\ge0} t(x)$, which is a contradiction. Therefore, $(x_n)_{n\ge1}$ is unbounded, and one can build a subsequence $(x_{nk})_{k\ge1}$ such that $\lim_{k\to\infty} x_{n_k} = +\infty$. We still have, of course, $\lim_{k\to\infty} t(x_{n_k}) = \sup_{x\ge0} t(x)$, but on the other hand, we have $\lim_{k\to\infty} t(x_{n_k}) = \lim_{x\to\infty} t(x)$. From Property 1 in the proof of Theorem 1, $\lim_{x\to\infty} t(x) = C$ and Property 5 follows.

Thus, if $\eta_c > C$, then $t(x)$ achieves its supremum at some finite positive real A. Let us consider the control $p_t = A/X_t$ for $X_t \ge A$. (Note that the value of the retransmission probability is left unspecified for $X_t < A$ because it does not affect the throughput.) Then from (4) $d_n = \lambda - t_n(A/n)$, and from Property 3 in the proof of Theorem 1 $\lim_{n\to\infty} d_n = \lambda - t(A)$. Then it follows from [21] that $(X_t)_{t>0}$ is ergodic if $\lambda < t(A)$ and from [13] and the Appendix that $(X_t)_{t\ge0}$ is nonergodic if $\lambda > t(A)$. Thus, the maximum stable throughput of the system is $t(A) = \sup_{x\ge0} t(x) = \eta_c$. □

Note that the closed-loop throughput obtained in Theorems 1 and 2 can be interpreted as $\eta_c = \sup_{N\sim P(x), x>0} E[C_N]$, that is as the supremum over x of the expected value of C_N if N is a Poisson distributed random variable with mean x. Note that if we were to follow the popular approximation [1], [2], [10], [16], [18], [24], [26] that assumes that the number of transmissions in each slot, N, is Poisson distributed, and if we could choose any positive number as the mean of N by regulating the retransmission probability, the throughput would be equal to the average number of successes per slot, $E[C_N]$, maximized over the mean of N. As in the usual collision channel, a wrong analysis leads to a correct conclusion. Several examples are gathered in Table I (see [8] for details).

Probably the most important conclusion of this section is that in general it is not necessary to compute the exact value of p_n^*, which would require a large amount of on-line computations, and seriously hinder any application of Theorem 1 to the case where the backlog is unknown. Two cases may occur. If $t(x)$ does not attain its supremum, from Property 5 in the proof of Theorem 2, we have $\eta_c = \eta_0 = C$ (e.g., this happens in the model developed in [6] for mobile users with pairwise transmissions). In this case no throughput improvement can be achieved by varying the retransmission probability, and therefore it is enough to restrict attention to the open-loop strategy studied in [8]. On the other

TABLE I
OPEN-LOOP AND CLOSED-LOOP THROUGHPUTS FOR SEVERAL MULTIPACKET CHANNELS

	C_n	$\eta_0 = \lim_{n\to\infty} C_n$	$\eta_c = \sup_{A>0} e^{-A} \sum_{n=1}^{\infty} C_n \frac{A^n}{n!}$
conventional collision channel	$1 \quad n=1$ $0 \quad n>1$	0	e^{-1}
q-frequency frequency hopping [6]	$n\left(1-\frac{1}{q}\right)^{n-1}$	0	$q\, e^{-1}$
mobile users with pairwise transmission [6]	1	1	1
capture - power discrimination [6]	$1 \quad n=1$ $\frac{1}{\beta^2} \quad n>1$	$\frac{1}{\beta^2}$	$\frac{1}{\beta^2}+\left(1-\frac{1}{\beta^2}\right)\exp\left(-\frac{\beta^2}{\beta^2-1}\right)$
capture - timing discrimination [3]	$1 \quad n=1$ $(1-Q)^n \quad n>1$	0	$\max_{A>0}\{(AQ-1)e^{-A}+e^{-AQ}\}$

hand, if there exists A, $0 < A < +\infty$, such that $t(A) = \sup_{x\geq 0} t(x)$, then we have shown in the proof of Theorem 2 that the control $p_t = A/X_t$ for $X_t \geq A$ yields a maximum stable throughput $t(A) = \eta_c$, meaning that the system is optimal. Hence, only A has to be computed, and this can be done before starting the operation of the system.

Although in most practical applications $(C_n)_n \geq 1$ does have a limit, it is worth noticing that Theorem 1 can be generalized to the case where C does not exist. It can be shown [9] that if the drift is minimized at each step, then the system is stable for $\lambda < \sup_{x\geq 0} t(x)$ and unstable for $\lambda > \sup_{x\geq 0} t(x) + \lim_{n\to\infty} \sup C_n - \lim_{n\to\infty} \inf C_n$. As in the open-loop system when $(C_n)_{n\geq 1}$ does not have a limit, nothing more can be said about the throughput without further information on the sequence $(C_n)_{n\geq 1}$. But the main drawback in such a case is that there may not exist any control $p_n = A/X_n$ that yields the optimal throughput.

The access rule for new packets that we have been considering so far is usually referred to as delayed first transmission (DFT). With this access rule, newly arrived packets are treated exactly in the same way as backlogged packets. Let us now examine what happens when on the contrary an immediate first transmission (IFT) rule is used, that is when new packets are transmitted with probability one in the slot immediately following their arrival. It has been proved in [8] that the open-loop throughput is the same for both first transmission rules. The closed-loop throughput on the other hand depends on the access rule. For instance, it is well known [4] that for the usual collision channel in the IFT case, the optimal retransmission probability is $p_n^* = \lambda_0 - \lambda_1/\lambda_0 n - \lambda_1$, yielding an optimal throughput $\lambda_0 e^{\lambda_1/\lambda_0} e^{-1}$, in contrast to the throughput $\eta_c = e^{-1}$ for the DFT case. In the multipacket channel with the IFT rule, the optimal throughput depends not only on the mean but on the whole distribution of new packet arrivals. Interestingly enough, it can be proved that both throughputs coincide when the new packet arrivals are Poisson distributed. Still with the same method as in the proof of Theorem 1, it can be easily shown that there exists a retransmission probability that minimizes the drift d_n at state n. With such a retransmission probability, the system with IFT rule is stable for $\lambda < \sup_{x\geq 0} T(x)$ and unstable for $\lambda > \sup_{x\geq 0} T(x)$, with $T(x) = e^{-x} \sum_{n=0}^{\infty} x^n/n! \sum_{j=0}^{\infty} \lambda_j C_{n+j}$, where we have defined $C_0 = 0$ for notational convenience. It can also be proved that a control of the form $p_n = A/X_n$ yields a maximum stable throughput $T(A)$. Since $\sup_{x\geq 0} T(x)$ depends on the whole new packet arrival distribution $(\lambda_n)_{n\geq 0}$, this result is not as conclusive as in the DFT case. This is because the stability region $\lambda < \sup_{x\geq 0} T(x)$ is actually given in the form of an implicit equation in λ, which cannot be solved in general without further specifications on the distribution $(\lambda_n)_{n\geq 0}$. For instance, this stability region could be empty. Consider, for example, the usual collision channel with possibly some added noise $0 < C_1 \leq 1$ and $C_n = 0$ for $n \geq 2$. Then $T(x) = C_1 e^{-x}(\lambda_1 + \lambda_0 x)$ and $T'(x) = C_1 e^{-x}(\lambda_0 - \lambda_1 - \lambda_0 x)$. Therefore, for any distribution such that $\lambda_0 < \lambda_1$, $T(x)$ is maximum at $T(0) + C_1\lambda_1$, and the stability region is empty since $C_1\lambda_1 \leq \lambda_1 \leq \lambda$. Note that in this sense, the immediate first transmission does not perform as well as the delayed first transmission with which the system can always be stabilized.

If there are solutions to $\lambda < \sup_{x\geq 0} T(x)$, then the best throughput achievable by the class of algorithms in (1) is $\nu_c = \sup\{\lambda : \lambda < \sup_{x\geq 0} T(x)\}$. This is what happens, for instance, when the new packet arrivals are Poisson distributed.

Theorem 3: If the new packet arrivals are Poisson distributed, the best throughput achievable with an IFT rule is the same as in the DFT case, $\nu_c = \sup_{x\geq 0} t(x)$.

Proof of Theorem 3: If $\lim_{n\to\infty} C_n = +\infty$, then $\eta_c = \nu_c = +\infty$. Assume now that $C < +\infty$. We get

$$T(x) = e^{-(x+\lambda)} \sum_{n=0}^{\infty} \frac{x^n}{n!} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} C_{n+k}$$

$$= e^{-(x+\lambda)} \sum_{n=1}^{\infty} \frac{C_n}{n!} (x+\lambda)^n. \tag{13}$$

Thus, in this case, $T(x)$ depends only on λ, and to clarify the proof below, we denote it by $T_\lambda(x)$

$$T_\lambda(x) = t(x+\lambda). \tag{14}$$

Assume that $t(x)$ does not achieve its supremum. Then from Property 5 in the proof of Theorem 2, we have $\eta_c = C = \lim_{x\to\infty} t(x)$ It follows from (14) that for any $\lambda > 0$, $\lim_{x\to\infty} T_\lambda(x) = C$. Therefore, for all $\lambda > 0$, $\sup_{x\geq 0} T_\lambda(x) \geq C$. Hence, for all $\lambda > 0$, $\sup_{x\geq 0} T_\lambda(x) = \sup_{x\geq 0} t(x)$, and by definition of ν_c, we finally get $\nu_c = \sup_{x\geq 0} t(x)$. Note that T_λ does not achieve its supremum, in the sense that if there existed $\lambda \in (0, \nu_c)$ and $x_\lambda \geq 0$ such that $\nu_c = T_\lambda(x_\lambda)$, we would have $\sup_{x\geq 0} t(x) = t(\lambda + x_\lambda)$.

Assume now that $t(x)$ does achieve its supremum. there exists $x_0 \geq 0$ such that $\sup_{x\geq 0} t(x) = t(x_0)$. Then for all λ in $[0, x_0]$. $T_\lambda(x_0 - \lambda) = \sup_{x\geq 0} t(x) \geq \sup_{x\geq 0} T_\lambda(x)$. Thus, for all $\lambda \in [0, x_0]$

$$\sup_{x\geq 0} T_\lambda(x) = \sup_{x\geq 0} t(x) = T_\lambda(x_0 - \lambda). \tag{15}$$

We have for all $x \geq 0$ $t(x) \leq x$, therefore $\sup_{x\geq 0} t(x) \leq x_0$. Together with (15), it follows that for all $\lambda \in (0, \sup_{x\geq 0} t(x))$, $\lambda < \sup_{x\geq 0} T_\lambda(x)$, and therefore $\nu_c \geq \sup_{x\geq 0} t(x) = \eta_c$. Since from (14) $\sup_{x\geq 0} T_\lambda(x) \leq \sup_{x\geq 0} t(x) = \eta_c$ for all λ, we get $\nu_c \leq \eta_c$ and finally $\nu_c = \eta_c = \sup_{x\geq 0} t(x)$. Note that from (14), T_λ reaches its supremum too, since for all $\lambda < \nu_c$, there exists $x_\lambda \geq 0$ such that $T_\lambda(x_\lambda) = \nu_c$.

Note that we have also shown in this proof that $T(x)$ reaches its supremum iff $t(x)$ does, which means that η_c can be achieved with a control of the form $p_n = A/X_n$ iff ν_c can. □

III. OPTIMAL CONTROL FOR THE MULTIPACKET CHANNEL

It is assumed from now on that the users do not have access to the value of the backlog, so the problem becomes one of control of the Markov chain with partial state information provided by the channel feedback. We build a backlog estimate S_t with feedback which is such that $Z_t = 0$ if slot t was empty, and $Z_t = \bar{0}$ otherwise. The results of the previous section strongly suggest that we should use as a retransmission probability $p_t = A/S_t$, where A is a point at which $t(x)$ achieves its supremum (according to Property 5, A is assumed to be finite). We show that the resulting control algorithm achieves the optimal maximum stable throughput η_c. This holds provided that the following assumption on the reception matrix is verified.

C0: There exists $\theta > 0$ and B such that for all $n \geq 1$, $\Sigma_{k=1}^{n} e^{\theta k} \epsilon_{nk} \leq B$.

The purpose of condition C0 is to bound the probability of having large numbers of simultaneous successes. Unbounded numbers of successes per slot are difficult to deal with because they may result in very large instantaneous errors in the backlog estimate. Note that condition C0 is likely to hold in most multipoint-to-point channels because of practical limitations on the receiver capabilities, and that it is verified for all the examples in Table I.

Theorem 4: Assume that there exists $A \in (0, +\infty)$ such that $t(A) = \sup_{x \geq 0} t(x)$, that the new packet arrivals $(A_t)_{t \geq 0}$ are exponential type[1], and that condition C0 holds. If $\alpha < 0$ and $\beta < 0$ verify the following two conditions[2]:

C1: $\beta > \lambda$

C2: $\beta(1 - e^{-A}) + \eta_c - \lambda + \alpha e^{-A} = 0$

then the control algorithm (cf. the control laws proposed in [15], [19], and [25])

$$p_t = \frac{A}{S_t}$$

$$S_{t+1} = \max\{A, S_t + \alpha I(Z_t = 0) + \beta I(Z_t = \bar{0})\}$$

has maximum stable throughput equal to η_c.

Proof of Theorem 4: The proof is based on the method developed in [30]. The idea is to use the properties of the homogeneous two-dimensional vector Markov chain of the backlog and its estimate $M_t = (X_t, S_t)$ to build a Lyapunov function whose drift is negative in the first quadrant of the (n, s) plane when $\lambda < \eta_c$. It turns out that this fails to hold in two cones of the state space, but it can be proved that the J-step drift of the Lyapunov function is negative for some integer J, and that this is enough to ensure that M_t is geometrically ergodic. It follows from Theorem 2 that M_t is nonergodic if $\lambda > \eta_c$. For substantial portions of the proof, the reader is referred to [9] because of space limitations.

Denote by $\tilde{X}_t = S_t - X_t$ the error in the backlog estimate. The first part of the proof mainly consists of computing and approximating the drifts of X_t and $\tilde{X}_t$ which are the basic building blocks for the Lyapunov function.

Denote by $c(n, s) = E[X_{t+1} - X_t | M_t = (n, s)]$ the backlog drift at state (n, s), and by $d(n, s) = E[\tilde{X}_{t+1} - \tilde{X}_t | M_t = (n, s)]$ the drift of the backlog error. For technical reasons, what we most often use in the proof are the truncated drifts, which correspond to the value of the drifts restricted to those paths where the variation in the backlog is bounded by some integer J, that is $c(n, s, J) = E[(X_{t+1} - X_t) I(|X_{t+1} - X_t| \leq J) | M_t = (n, s)]$ and $d(n, s, J) = E[(\tilde{X}_{t+1} - \tilde{X}_t) I(|X_{t+1} - X_t| \leq J) | M_t = (n, s)]$. Clearly, these truncated drifts will be good approximations of $c(n, s)$ and $d(n, s)$, respectively, when J is large. It will turn out that the drifts depend primarily on the ratio $x = n/s$ for large values of n or s. Thus, it is convenient to define the following two regions in the (n, s) plane:

$$C(\lambda_0, \lambda_1) = \{(n, s) : n \geq 0, s \geq 0, 1 + \lambda_0 \leq \frac{n}{s} \leq 1 + \lambda_1\}$$

$$U_M = \{(n, s) : n \geq M \text{ or } s \geq M\}$$

where λ_0 and λ_1 are such that $-\infty \leq \lambda_0 \leq \lambda_1 \leq +\infty$. The aim of the first part of the proof is to show Proposition 1 below which summarizes all the properties of the drifts that are needed for our purposes (see Fig. 1).

[1] A_t is exponential type if there exists $d > 0$ such that $E[e^{dA_t}]$ is finite. For instance, this is true if A_t is Poisson distributed.

[2] Conditions C1 and C2 define half a straight line in the plane, and therefore an infinite number of possible estimation schemes, all of them yielding the same throughput.

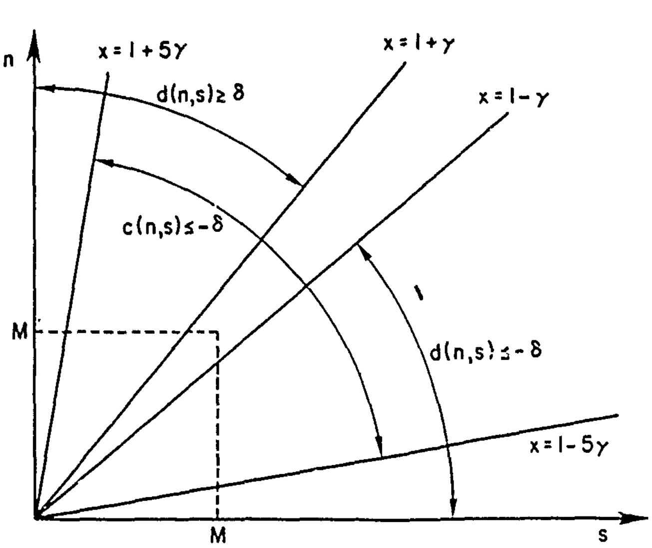


Fig. 1. Drift properties (Proposition 1).

Proposition 1: There exist $\gamma \in (0, 1/5)$, $\delta > 0$, and an integer $J_0 > 0$ such that for all $J \geq J_0$:

i) for all $(n, s) \in C(-5\gamma, 5\gamma) \cap U_M$, $c(n, s) \leq -\delta$ and $c(n, s, J) \leq -\delta + \nu(J)$;

ii) for all $(n, s) \in C(-\infty, -\gamma) \cap U_M$, $d(n, s) \leq -\delta$ and $d(n, s, J) \leq -\delta + \nu(J)$;

iii) for all $(n, s) \in C(\gamma, +\infty) \cap U_M$, $d(n, s) \geq \delta$ and $d(n, s, J) \geq \delta - \nu(J)$

where $\nu(J)$ is a nonnegative function which goes to zero a J goes to infinity.

The detailed proof of Proposition 1 can be found in [9]. After computing the value of the drifts

$$c(0, s) = \lambda \tag{16a}$$

$$c(n, s) = \lambda - \sum_{j=1}^{n} \binom{n}{j} \left(\frac{A}{s}\right)^j \left(1 - \frac{A}{s}\right)^{n-j} C_j \quad (n \geq 1) \tag{16b}$$

$$d(0, s) = \max\{A - s, \alpha\} - \lambda \tag{17a}$$

$$d(n, s) = \beta - \lambda + (\max\{A - s, \alpha\} - \beta) \left(1 - \frac{A}{s}\right)^n + \sum_{j=1}^{n} \binom{n}{j} \left(\frac{A}{s}\right)^j \left(1 - \frac{A}{s}\right)^{n-j} C_j \quad (n \geq 1) \tag{17b}$$

we work out upper and lower bounds by truncating the sums (16) and (17) to a fixed number of terms, and then we approximate those bounds as a function of the sole variable n/s. The main idea is that the dynamic behavior of the Markov vector $M_t = (X_t, S_t)$ depends essentially on the ratio X_t/S_t. For instance, if x is nearly equal to 1, the backlog estimate is close to its ideal value, and we should have $c(n, s) < 0$ since the backlog drift is negative in the perfect state information case. Also, a well-behaved estimate should be such that if $x < 1$, then the error $s - n$ is positive, and therefore should have a negative drift $d(n, s) < 0$ (see [15]). In the same way, we expect to have $d(n, s) > 0$ for $x > 1$.

Let us define the following Lyapunov function:

$$V(n, s) = \max\left\{n, \frac{1 + 3\gamma}{3\gamma}(n - s), \frac{1 - 3\gamma}{3\gamma}(s - n)\right\}$$

where the constants have been chosen so that V is continuous. $V(n, s)$ is equal to the first, second, and third term inside the bracket when (n, s) is in $C(-3\gamma, 3\gamma)$, $C(3\gamma, +\infty)$, and $C(-\infty, -3\gamma)$, respectively. Notice that V is defined so as to take the best advantage of the drift properties listed in Proposition 1. For instance, when $V(n, s)$ is equal to n, then the Markov chain M_t belongs to $C(-3\gamma, 3\gamma)$ which is included in $C(-5\gamma, 5\gamma)$ where the backlog drift is negative provided that either n or s is sufficiently large. Similar comments can be made about the other two regions. Unfortunately, this does not enable us to conclude that the drift of the Lyapunov function is negative in U_M because M_{t+1} may well be in a different region than M_t. However, this change of region becomes unlikely if we exclude a small zone around the lines $x = 1 \pm 3\gamma$ where V changes definition and indeed the second part of this proof consists of showing that the Lyapunov function has a negative drift in the remainder of the state space.

Proposition 2: There exist $M_0 \geq 0$ and $\delta_0 > 0$ such that for all $N \geq M_0$ and for all $(n, s) \in U_N \cap [C(-\infty, -4\gamma) \cup C(-2\gamma, 2\gamma) \cup C(4\gamma, \infty)]$,

$$E[V(M_{t+1}) - V(M_t) | M_t = (n, s)] < -\delta_0.$$

Proof of Proposition 2: We consider separately likely and unlikely events

$$\begin{aligned} E[V(M_{t+1}) - V(M_t) | M_t = (n, s)] &= E[(V(M_{t+1}) - V(M_t)) I(|A_t - \Sigma_t| \leq J) | M_t = (n, s)] \\ &\quad + E[(V(M_{t+1}) - V(M_t)) I(|A_t - \Sigma_t| > J) | M_t = (n, s)]. \end{aligned} \tag{18}$$

We start by showing that the first term, which corresponds to likely events, is negative when J is large by using the properties of the truncated drifts from Proposition 1 and a simple geometric result. The lemma below, whose proof is in [9], gives a measure of how much a cone $C(\lambda_0, \lambda_1)$ expands if each of its points is allowed to move of some distance that cannot exceed B in absolute value along each axis.

Lemma: Consider $\gamma > 0$, $B > 0$, and $\gamma - 1 < \lambda_0 < \lambda_1 < +\infty$; and assume that $|n - n'| \leq B$, $|s - s'| \leq B$, and $Q \geq B/\gamma(1 + |\lambda_1|)(\lambda_1 + 2 + \gamma)$. Then:

1) $(n, s) \in C(\lambda_0, \infty) \cap U_Q$
$$\Rightarrow (n', s') \in C(\lambda_0 - \gamma, \infty) \cap U_{Q-B}$$

2) $(n, s) \in C(-\infty, \lambda_1) \cap U_Q$
$$\Rightarrow (n', s') \in C(-\infty, \lambda_1 + \gamma) \cap U_{Q-B}$$

3) $(n, s) \in C(\lambda_0, \lambda_1) \cap U_Q$
$$\Rightarrow (n', s') \in C(\lambda_0 - \gamma, \lambda_1 + \gamma) \cap U_{Q-B}.$$

Set $B(J) = \max\{J, |\alpha| + \beta\}$, and define $Q(J)$ to be any real such that $Q(J) \geq \max\{B(J) + M, B(J)/\gamma(1 + 4\gamma)(2 + 3\gamma)\}$. We have $|S_{t+1} - S_t| \leq |\alpha| + \beta \leq B(J)$, and if $|A_t - \Sigma_t| \leq J$, then $|X_{t+1} - X_t| \leq J \leq B(J)$. From the lemma, $Q(J)$ is such that

$$M_t \in C(-2\gamma, 2\gamma) \cap U_{Q(J)} \Rightarrow M_{t+1} \in C(-3\gamma, 3\gamma) \cap U_M \tag{19}$$

$$M_t \in C(4\gamma, \infty) \cap U_{Q(J)} \Rightarrow M_{t+1} \in C(3\gamma, \infty) \cap U_M \tag{20}$$

$$M_t \in C(-\infty, -4\gamma) \cap U_{Q(J)} \Rightarrow M_{t+1} \in C(-\infty, -3\gamma) \cap U_M \tag{21}$$

where M has been defined in Proposition 1. Assume, for instance that M_t belongs to $C(-2\gamma, 2\gamma) \cap U_{Q(J)}$. From (19), $M_{t+1} \in C(-3\gamma, 3\gamma) \cap U_M \cap C(-5\gamma, 5\gamma) \cap U_M$. Hence, if $J \geq J_0$, we can apply Proposition 1 i):

$$E[(V(M_{t+1}) - V(M_t)) I(|A_t - \Sigma_t| \leq J) | M_t = (n, s)] = c(n, s, J) \leq -\delta + \nu(J).$$

If M_t belongs to the other two regions, $C(4\gamma, \infty) \cap U_{Q(J)}$ or $C(-\infty, -4\gamma) \cap U_{Q(J)}$, a similar argument holds, using Proposition 1 iii) and ii), respectively, along with (20) and (21). It follows that for all $J \geq J_0$ and for all $(n, s) \in U_{Q(J)} \cap [C(-\infty, -4\gamma) \cup C(-2\gamma, 2\gamma) \cup C(4\gamma, \infty)]$

$$E[(V(M_{t+1}) - V(M_t)) I(|A_t - \Sigma_t| \leq J) | M_t = (n, s)] \leq -\delta_1 + \nu_1(J) \tag{22}$$

with $\delta_1 = \min\{1, 1 - 3\gamma/3\gamma\}\delta$ and $\nu_1(J) = \nu(J)\ 1 + 3\gamma/3\gamma$.

To deal with the second term on the right-hand side of (18), we consider the further decomposition

$$\begin{aligned} E[(V(M_{t+1}) - V(M_t)) I(|A_t - \Sigma_t| > J) | M_t = (n, s)] &= E[(V(M_{t+1}) - V(M_t)) I(A_t > \Sigma_t + J) | M_t = (n, s)] \\ &\quad + E[(V(M_{t+1}) - V(M_t)) I(\Sigma_t > A_t + J) | M_t = (n, s)]. \end{aligned} \tag{23}$$

Let us denote by $T_1(n, s, J)$ and $T_2(n, s, J)$ the two terms on the right-hand side of (23). The first term $T_1(n, s, J)$ corresponds to a case where the variation in the backlog is bounded below, and can be shown to vanish as J increases by using the sole fact that the mean arrival rate λ is finite. Consider now $T_2(n, s, J)$. If $M_t = (n, s)$ belongs to a region such that $x = n/s > x_0$, then x_0 can be chosen large enough so that if M_{t+1} belongs to $C(-\infty, -3\gamma)$, then the error in the backlog estimate which results from the large number of successes just compensates the initial error $n - s \gg 0$. On the other hand, when M_t belongs to any region such that x is bounded above, then $E[\Sigma_t I(\Sigma_t > J) | M_t = (n, s)]$ goes to zero uniformly in (n, s) and $T_2(n, s, J)$ can be dealt with by using the following rather crude bound for the variation of V:

$$|V(M_{t+1}) - V(M_t)| \leq \max\left\{1, \frac{1+3\gamma}{3\gamma}, \frac{1-3\gamma}{3\gamma}\right\} \cdot (|\alpha| + \beta + |A_t - \Sigma_t|) \leq R(1 + |A_t - \Sigma_t|) \tag{24}$$

where R is some positive constant. It is shown in [9] that

$$E[(V(M_{t+1}) - V(M_t)) I(|A_t - \Sigma_t| > J) | M_t = (n, s)] \leq \nu_2(J) + \epsilon_J(n, s) \tag{25}$$

where $\lim_{J \to \infty} \nu_2(J) = 0$, and $\epsilon_J(n, s)$ is a nonnegative function that depends on J, and goes to zero as either n or s goes to infinity.

By using (22), (25), and the decomposition (18), we get the desired result that the drift of V is negative in this part of the state space: fix an integer $J_{\min}$ such that $J_{\min} \geq J_0$ and that for all $J \geq J_{\min}$, $\nu_1(J) + \nu_2(J) \leq \delta_1/3$. Then from (22) and (25), we have for all $(n, s) \in U_{Q(J_{\min})} \cap [C(-\infty, -4\gamma) \cup C(-2\gamma, 2\gamma) \cup C(4\gamma, \infty)]$,

$$E[V(M_{t+1}) - V(M_t) | M_t = (n, s)] \leq -\frac{2}{3}\delta_1 + \epsilon_{J_{\min}}(n, s).$$

Then we can choose an $M_0 > Q(J_{\min})$ which is large enough so that $\epsilon_{J_{\min}}(n, s) < \delta_1/3$ for all (n, s) in U_{M_0}. □

This concludes the second part of the proof. Unfortunately, it is not always true that the drift of V is negative outside a finite subset of the state space. For instance, we have proved that in the case of the usual collision channel with Poisson new packet arrivals, there exist constants $B_{ex} > 0$ and M_{ex} such that for all $(n, s) \in U_{M_{ex}}$ for which $x = 1 \pm 3\gamma$, and for all α and β verifying C1 and C2, $E[V(M_{t+1}) - V(M_t) | M_t = (n, s)] > B_{ex}$. However, discontinuities around the lines $x = 1 \pm 3\gamma$ cancel out when one waits long enough, and in the last part of this proof we show that the J-step drift of V, $E[V(M_{t+J}) - V(M_t) | M_t = (n, s)]$ is negative for some integer J.

Proposition 3: There exist $J_f > 0$, $\rho > 0$, and $M_f > 0$ such that for all $(n, s) \in U_{M_f}$

$$E[V(M_{t+J_f}) - V(M_t) | M_t = (n, s)] \le -\rho.$$

Proof of Proposition: One of the main problems in dealing with the J-step drift of V is to control the changes of regions between M_t and M_{t+J}. To this end, we define the stopping time

$$\tau_J = \min\left\{s \ge 0, \left|\sum_{k=0}^{s} (A_{t+k} - \Sigma_{t+k})\right| > J^3\right\}.$$

If $\tau_J \ge J$, then for $1 \le k \le J$, $|X_{t+k} - X_t| \le J^3$ and $|S_{t+k} - S_t| \le J(|\alpha| + \beta)$. Thus, if we define $B'(J) = \max\{J(|\alpha| + \beta), J^3\}$, and $Q'(J)$ to be any integer such that $Q'(J) \ge B'(J) + \max\{M_0, M\}$ and $Q'(J) \ge 2B'(J)/\gamma(1 + 9/2\gamma)(5\gamma + 2)$, then, still assuming that $\tau_J \ge J$, we get from the lemma for $0 \le k \le J$

$$M_t \in C\left(-\infty, -4\gamma - \frac{\gamma}{2}\right) \cap U_{Q'(J)} \Rightarrow M_{t+k} \in C(-\infty, -4\gamma) \cap U_{M_0} \quad (26)$$

$$M_t \in C\left(-2\gamma + \frac{\gamma}{2}, 2\gamma - \frac{\gamma}{2}\right) \cap U_{Q'(J)} \Rightarrow M_{t+k} \in C(-2\gamma, 2\gamma) \cap U_{M_0} \quad (27)$$

$$M_t \in C\left(4\gamma + \frac{\gamma}{2}, \infty\right) \cap U_{Q'(J)} \Rightarrow M_{t+k} \in C(4\gamma, \infty) \cap U_{M_0} \quad (28)$$

$$M_t \in C\left(-4\gamma - \frac{\gamma}{2}, -2\gamma + \frac{\gamma}{2}\right) \cap U_{Q'(J)} \Rightarrow M_{t+k} \in C(-5\gamma, -\gamma) \cap U_M \quad (29)$$

$$M_t \in C\left(2\gamma - \frac{\gamma}{2}, 4\gamma + \frac{\gamma}{2}\right) \cap U_{Q'(J)} \Rightarrow M_{t+k} \in C(\gamma, 5\gamma) \cap U_M. \quad (30)$$

In other words, we have partitioned the plane into two zones

$$Z_N = C\left(-\infty, -4\gamma - \frac{\gamma}{2}\right) \cup C\left(-2\gamma + \frac{\gamma}{2}, 2\gamma - \frac{\gamma}{2}\right) \cup C\left(4\gamma + \frac{\gamma}{2}, \infty\right),$$

and

$$Z_P = C\left(-4\gamma - \frac{\gamma}{2}, -2\gamma + \frac{\gamma}{2}\right) \cup C\left(2\gamma - \frac{\gamma}{2}, 4\gamma + \frac{\gamma}{2}\right).$$

Then we have chosen $Q'(J)$ such that if M_t belongs to Z_N which is slightly smaller than the region in which the drift of the Lyapunov function is negative, and if $\tau_J \ge J$, then the Markov chain remains in the region in which Proposition 2 applies up to time $t + J$ (see (26)–(28) and Fig. 2). $Q'(J)$ is also such that if M_t is in Z_P and if $\tau_J \ge J$, then up to time $t + J$ the chain stays in a region such that two out of the three properties of Proposition 1 hold at each step (see (29), (30), and Fig. 3).

We start by showing that the J-step drift of V is negative at (n, s) when (n, s) belongs to Z_N. We decompose the J-step drift of V

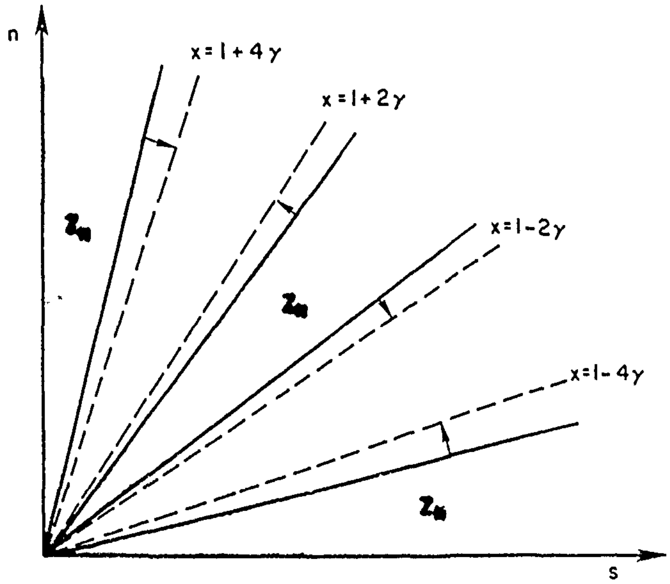


Fig. 2. If $M_t \in Z_N \cap U_{Q'(J)}$ and if $\tau_J \ge J$, then M_{t+1} belongs to the region where the drift of V is negative.

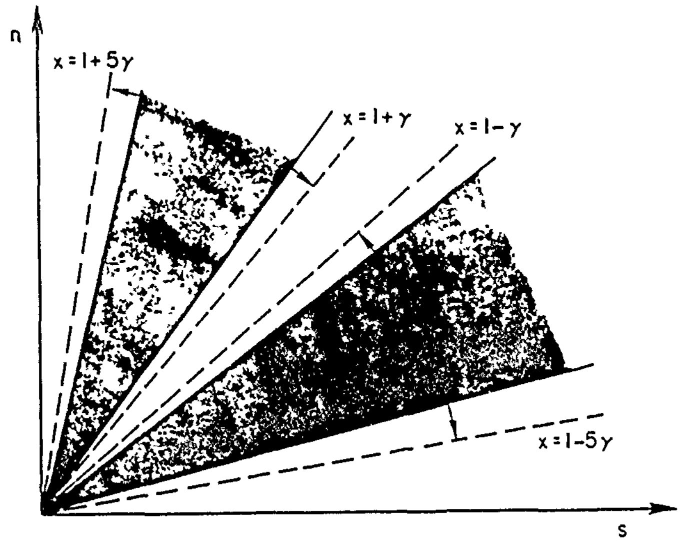


Fig. 3. If $M_t \in Z_P \cap U_{Q'(J)}$ and if $\tau_J \ge J$, then M_{t+1} belongs to a region where two properties of Proposition 1 hold.

as follows:

$$E[V(M_{t+J}) - V(M_t) | M_t = (n, s)] = \sum_{k=0}^{J-1} E[E[V(M_{t+k+1}) - V(M_{t+k}) | M_{t+k}] \cdot I(\tau_J \ge J) | M_t = (n, s)] + \sum_{k=0}^{J-1} E[E[V(M_{t+k+1}) - V(M_{t+k}) | M_{t+k}] I(\tau_J < J) | M_t = (n, s)]. \quad (31)$$

Denote by $U_1(J, n, s)$ and $U_2(J, n, s)$ the two sums on the right-hand side of (31). If $\tau_J \ge J$, then (26)–(28) hold, and therefore we can apply Proposition 2

$$U_1(J, n, s) \le -J\delta_0 P[\tau_J \ge J | M_t = (n, s)]. \quad (32)$$

Let us now show that $\tau_J < J$ is indeed an unlikely event, the probability of which goes to zero as $1/J$ uniformly in (n, s)

$$P[\tau_J<J|M_t=(n, s)]$$
$$\leq \sum_{k=0}^{J-1} P\left[\left|\sum_{l=0}^{k}(A_{t+l}-\Sigma_{t+l})\right|>J^3|M_t=(n, s)\right]$$
$$\leq \sum_{k=0}^{J-1} P\left[\sum_{l=0}^{k}A_{t+l}>J^3\right]$$
$$+\sum_{k=0}^{J-1} P\left[\sum_{l=0}^{k}\Sigma_{t+l}>J^3|M_t=(n, s)\right].$$

From Markov's inequality we have

$$P[\tau_J<J|M_t=(n, s)]\leq\frac{1}{J^3}\sum_{k=0}^{J-1}(k+1)\lambda$$
$$+\frac{1}{J^3}\sum_{k=0}^{J-1}\sum_{l=0}^{k}E[\Sigma_{t+l}|M_t=(n, s)].$$

Denoting by B_η an upper bound on the sequence $t_n(p_n^*)$, it follows from Section II that $E[\Sigma_{t+l}|M_t = (n, s)] = E[E[\Sigma_{t+l}|M_{t+1}]|M_t = (n, s)] \leq B_\eta$, so we get

$$P[\tau_J<J|M_t=(n, s)]\leq\frac{\lambda+B_\eta}{2}\frac{J+1}{J^2}\leq\frac{B_\tau}{J} \tag{33}$$

where B_τ is some positive constant. From (24), it is easy to check that the drift of V is bounded by some positive constant B_V, so that

$$U_2(J, n, s)\leq JB_VP[\tau_J<J|M_t=(n, s)]. \tag{34}$$

Considering (31), (32), (33), and (34), we get

$$E[V(M_{t+J})-V(M_t)|M_t=(n, s)]\leq-\delta_0J+(B_V+\delta_0)B_\tau.$$

Therefore, there exist constants $\mu_1 > 0$ and $J_1 > 0$ such that for all $J \geq J_1$ and for all $(n, s) \in U_{Q'(J)} \cap Z_N$,

$$E[V(M_{t+J})-V(M_t)|M_t=(n, s)]\leq-J\mu_1. \tag{35}$$

We now proceed to show that the J-step drift of the Lyapunov function is negative in the remaining part of the state space Z_P consisting of the two cones around $x = 1 \pm 3\gamma$. This is done in two steps. We first show that the J-step drift of V restricted to likely events $\{\tau_J \geq J\}$ goes to $-\infty$ as J increases, and then we prove that the J-step drift of V restricted to unlikely events $\{\tau_J < J\}$ is bounded above independent of J.

Assume, for instance, that $(n, s) \in C(\gamma - \gamma/2, 4\gamma + \gamma/2) \cap E_{Q'(J)}$. The difficulty here is that V can take two possible values, and therefore Proposition 1 cannot be used directly. If $\tau_J \geq J$, then from (30) $M_{t+k} \in C(\gamma, 5\gamma) \cap U_M$ for $0 \leq k \leq J$, so that $V(M_{t+k}) = \max\{X_{t+k}, (1 + 3\gamma)/3\gamma(X_{t+k} - S_{t+k})\}$. Therefore,

$$E[(V(M_{t+J})-V(M_t))I(\tau_J\geq J)|M_t=(n, s)]$$
$$=E\left[\max\left\{X_{t+J}, \frac{1+3\gamma}{3\gamma}(X_{t+J}-S_{t+J})\right\}\cdot I(\tau_J\geq J)|M_t=(n, s)\right]$$
$$-E\left[\max\left\{X_t, \frac{1+3\gamma}{3\gamma}(X_t-S_t)\right\}\cdot I(\tau_J\geq J)|M_t=(n, s)\right]$$
$$\leq E\left[\max\left\{X_{t+J}-X_t, \frac{1+3\gamma}{3\gamma}(-\tilde{X}_{t+J}+\tilde{X}_t)\right\}\cdot I(\tau_J\geq J)|M_t=(n, s)\right]$$

since $\max\{a, b\} - \max\{c, d\} \leq \max\{a - c, b - d\}$. Then using the fact that $\max\{a, b\} \leq \max\{0, a + f\} + \max\{0, b + f\} - f$ for $f \geq 0$, we get

$$E[(V(M_{t+J})-V(M_t))I(\tau_J\geq J)|M_t=(n, s)]$$
$$\leq E\left[\max\left\{0, X_{t+J}-X_t+\delta_1\frac{J}{2}\right\}\cdot I(\tau_J\geq J)|M_t=(n, s)\right]$$
$$+E\left[\max\left\{0, \frac{1+3\gamma}{3\gamma}(-\tilde{X}_{t+J}+\tilde{X}_t)+\delta_1\frac{J}{2}\right\}\cdot I(\tau_J\geq J)|M_t=(n, s)\right]$$
$$-E\left[\delta_1\frac{J}{2}I(\tau_J\geq J)|M_t=(n, s)\right] \tag{36}$$

where $\delta_1 = \min\{1, (1-3\gamma)/3\gamma\}$ has been defined in (22). We show that the first two terms on the right-hand side of (36) are bounded. Since (33) $\lim_{J\to\infty} - \delta_1J/2P[\tau_J \geq J] = -\infty$, this will be sufficient to prove that $\lim_{J\to\infty} E[(V(M_{t+J}) - V(M_t))I(\tau_J \geq J)|M_t = (n, s)] = -\infty$. Define $W_k = X_{t+k} - X_t + k\gamma_1/2$ and $\tilde{F}_k = F_{t+k}$; where F_t is the sigma-field generated by $\{A_s, s \leq t - 1; X_s, s \leq t\}$, representing the history of the process $(M_t)_{t\geq 0}$ up to time t. To prove that the first term in (36) is bounded, we show that there exists $\phi > 0$ such that $(Y_k, \tilde{F}_k)$ is a supermartingale, with $Y_k = e^{\phi W_k}I(\tau_J \geq k)$. We need to show that $E[Y_{k+1}|\tilde{F}_k] \leq Y_k$, which is equivalent to

$$E[e^{\phi(X_{t+k+1}-X_t+(k+1/2)\delta_1)}I(\tau_J\geq k+1)|F_{t+k}]$$
$$\leq e^{\phi(X_{t+k}-X_t+(k/2)\delta_1)}I(\tau_J\geq k)$$

since $I(\tau_J \geq k + 1) = I(\tau_J \geq k)I(T_J \geq k + 1)$, and $I(\tau_J \geq k)$ is measurable with respect to F_{t+k}

$$I(\tau_J\geq k)E[e^{\phi(X_{t+k+1}-X_{t+k}+\delta_1/2)}|F_{t+k}]\leq I(\tau_J\geq k). \tag{37}$$

Now if $\tau_J \geq k$, then from (30), $M_{t+k} \in C(\gamma, 5\gamma) \cap U_M$. Lemma 2.2 in [11] states that if X is a random variable such that $|X|$ is stochastically dominated by an exponential type random variable Z, and if the expectation of X is strictly negative, $E|X| < -\epsilon$, then there exist two constants $\eta > 0$ and $\rho < 1$ such that $E[e^{\eta X}] < \rho < 1$. Hence, there exists $\phi > 0$ such that

for all $(n, s) \in C(-5\gamma, 5\gamma) \cap U_M$,

$$E[e^{\phi(X_{t+1}-X_t+\delta/2)}|M_t=(n, s)]<1 \tag{38a}$$

for all $(n, s) \in C(-\infty, -\gamma) \cap U_M$,

$$E[e^{\phi(\tilde{X}_{t+1}-\tilde{X}_t+\delta/2)}|M_t=(n, s)]<1 \tag{38b}$$

for all $(n, s) \in C(\gamma, \infty) \cap U_M$,

$$E[e^{\phi(-\tilde{X}_{t+1}+\tilde{X}_t+\delta/2)}|M_t=(n, s)]<1. \tag{38c}$$

It follows from (37) and (38a) that $(Y_k, \tilde{F}_k)$ is a supermartingale. Therefore,

$$E[Y_J|\tilde{F}_0]=E[^{\phi W_J}I(\tau_J\geq J)|F_t]\leq E[Y_0|\tilde{F}_0]=1. \tag{39}$$

Finally, considering that $\max\{0, x\} \leq 1/\phi\, e^{\phi x}$, it follows from (39) that the first term in (36) is bounded. Using (30) and (38c), it can be shown with the same method that the second term in (36) is also bounded. Thus, threre exists a constant B_T independent of J

such that

$$E[(V(M_{t+J}-V(M_t))I(\tau_J\geq J)|M_t=(n,s)]\leq B_T-\frac{J}{2}\delta_1 P[\tau_J\geq J].$$

The case $(n, s) \in C(-4\gamma-\gamma/2, -2\gamma+\gamma/2)\cap U_{Q'(J)}$ can be dealt with in a similar way, using (38a) and (38b). Therefore, we have shown that there exist $\mu_2 > 0$ and $J_2 > 0$ such that for all $J \geq J_2$ and for all $(n, s) \in Y_{Q'(J)}\cap Z_P$

$$E[(V(M_{t+J})-V(M_t))I(\tau_J\geq J)|M_t=(n,s)]<-J\mu_2. \quad (40)$$

It is shown in [9] that there exist a constant $B > 0$, a function $\nu_3(J)$ with $\lim_{J\to\infty}\nu_3(J)$, and a nonnegative function $\nu_J(M)$ depending on J verifying $\lim_{M\to\infty}\nu_J(M) = 0$, such that for all $(n, s) \in U_{Q'(J)+M_1}\cap Z_P$,

$$E[(V(M_{t+J})-V(M_t))I(\tau_J<J)|M_t=(n,s)]$$
$$<B+\nu_3(J)+\nu_J(M_1). \quad (41)$$

We are now ready to conclude the proof of Proposition 3. From (40) and (41), we have for all $(n, s) \in U_{Q'(J)+M_1}\cap Z_P$, $E[V(M_{t+J}) - V(M_t)|M_t = (n, s)] \leq B - J\mu_2 + \nu_3(J) + \nu_J(M_1)$. Fix an integer $J_f \geq \max\{J_1, J_2\}$ such that for all $J \geq J_f$, $B - J\mu_2 + \nu\nu_3(J) < -\mu_2$. Then for all $(n, s) \in U_{Q'(J_f)+M_1}\cap Z_P$, we have $E[V(M_{t+J_f}) - V(M_t)|M_t = (n, s)] \leq -\mu_2 + \nu_{J_f}(M_1)$. On the other hand, we also have from (44), for all $(n, s) \in U_{Q'(J_f)+M_1}\cap Z_P$

$$E[V(M_{t+J_f})-V(M_t)|M_t=(n,s)]\leq-\mu_1 J_f.$$

Now fix M_1 large enough so that $\nu_J(M_1) \leq \mu_2/2$. Then define $M_f = Q'(J_f) + M_1$, and $\rho = \min\{\mu_2/2, J_f\mu_1\}$. □

We can now conclude that $(M_t)_{t\geq 0}$ is geometrically ergodic for $\lambda < \eta_c$ by invoking the following result.

Theorem (Hajek [11]): Let $\{W_t\}$ be a sequence of random variables adapted to an increasing family of σ-fields $\{F_t\}$. Suppose that W_0 is deterministic, that $\{W_t, F_t\}$ is exponential type, and that for some $\epsilon > 0$ and $a > 0$ we have $E[(W_{t+1} - W_t - \epsilon)I(W_t > a)|F_t] \leq 0$ for all $t \geq 0$. Then for each value of W_0 the stopping time $\tau = \min\{t \geq 0;\ W_t \leq a\}$ is exponential type.

Define $W_t = V(M_{tJ_f})$ and $a = M_f\max\{1, (1+3\gamma)/3\gamma, (1-3\gamma)/3\gamma\}$. If $V(M_t) > a$, then $M_t \in U_{M_f}$. From (24) and C0 $(V(M_t), F_t)$ is exponential type since A_t is. From Proposition 3, we can apply Hajek's result to our system to conclude that $\tau = \min\{t \geq 0,\ V(M_{tJ_f}) \leq a\}$ is exponential type for any initial state. Since $V(M_t) \leq a$ implies that $X_t \leq a$ and $S_t \leq a/(1-3\gamma)$, it follows that $\tau' = \min\{t \geq 0,\ X_{tJ_f} \leq a$, and $S_{tJ_f} \leq a/(1-3\gamma)\}$ is also exponential type for any initial state, as well as $\tau'' = \min\{t \geq 0, X_t \leq a$, and $S_t \leq a/(1-3\gamma)\}$. Hence, it follows from [14] that (X_t, S_t) is geometrically ergodic, concluding the proof of Theorem 4. □

IV. Stability Proof Via Mikhailov's Theorem

Mikhailov [19, Theorem 3] has recently found a powerful sufficient condition to guarantee the stability of a Markov process taking values on $R^+ \times R^+$. This result can be used to weaken the sufficient conditions we imposed in Section III and obtain a much more simple proof of stability. However, the form of stability used by Mikhailov is weaker than the geometric ergodicity used in Section III.

Let M_t be a discrete-time Markov process taking values in $Y \subseteq R^n$, $U(r) = \{x \in R^n: \|x\| \leq r\}$, and $\tau_x(S) = \min\{t \geq 0: M_t \in S|M_0 = x\}$, i.e., $\tau_x(S)$ is the time it takes to reach the set S from x. Then we say that the process M_t is stable if there exist constants c_1 and c_2 such that $E[\tau_x(U(r))] \leq c_1\|x\| + c_2$ for all $x \in Y$. Using this definition of stability we show the following result which is analogous to Theorem 4.

Theorem 5: Suppose that:

i) the number of new packet arrivals per slot has finite second moment $E[A_t^2] < +\infty$;

ii) there exists $A \in (0, +\infty)$ such that $t(A) = \sup_{x\geq 0} t(x)$;

iii) C0″: there exists $B < +\infty$ such that for all $n \geq 1$, $\Sigma_{k=1}^n k^2\epsilon_{nk} \leq B$.

Fix $\lambda < \eta_c$ and $\xi > 0$ such that $\lambda < t(A\xi)$. Choose $\alpha < 0$ and $\beta > 0$ such that

$$C1':\ \beta(e^{A\xi}-1)=\frac{\lambda-t(A\xi)}{\xi}e^{A\xi}-\alpha$$

$$C_2':\ \beta>m_\xi(\lambda)=\sup_{x>0,x\neq\xi}\frac{\lambda-t(Ax)-xe^{-A(x-\xi)}\dfrac{\lambda-t(A\xi)}{\xi}}{x-xe^{-A(x-\xi)}}.$$

Then the control algorithm

$$p_t=\frac{A}{S_t}$$

$$S_{t+1}=\max\{A, S_t+\alpha I(Z_t=0)+\beta I(Z_t=\bar{0})\}$$

is stable.

Proof of Theorem 5: Let us state first Mikhailov's Theorem (cf. [35] for an exposition of this result and its application in the decentralized control of the conventional collision channel).

Theorem (Mikhailov [19]): Let $M_t = (X_t, S_t)$ be a homogeneous Markov process on $R^+ \times R_0^+$, with drifts

$$(c(n,s), e(n,s))=E[M_{t+1}-M_t|M_t=(n,s)].$$

Suppose that:

i) there exists $B < +\infty$ such that for all $(n, s) \in R^+ \times R_0^+$, $E[\|M_{t+1} - M_t\|^2|M_t = (n, s)] \leq B$;

ii) for all $\psi \in (0, +\infty)$, the drifts $(c(n, n/\psi), e(n, n/\psi))$ converge uniformly in ψ as n goes to infinity to $(c(\psi), e(\psi))$;

iii) the limit drifts $(c(\psi), e(\psi))$ are differentiable on $[0, +\infty)$, with $(e(0), e(0)) = \lim_{s\to\infty}(c(0, s), e(0, s))$;

iv) there exists $\epsilon > 0$ such that if $c(\psi_0) = \psi_0 e(\psi_0)$, then $c(\psi_0) < -\epsilon$.

Then M_t is stable.

Since both the new packet arrivals and the rows of the reception matrix have finite variance, it is easy to check that condition i) in Mikhailov's Theorem holds

$$E[\|M_{t+1}-M_t\|^2|M_t=(n,s)]$$
$$=E[(X_{t+1}-X_t)^2+(S_{t+1}-S_t)^2|M_t=(n,s)].$$

Now $E[(S_{t+1} - S_t)^2|M_t = (n, s)] \leq \alpha^2 + \beta^2$, and from (2)

$$E[(X_{t+1}-X_t)^2|M_t=(n,s)]\leq E[A_t^2]+E[\Sigma_t^2|M_t=(n,s)].$$

From C0′ the variance of the number of successes is also bounded

$$E[\Sigma_t^2|M_t=(n,s)]$$
$$=\sum_{k=1}^n k^2\sum_{j=k}^n\binom{n}{j}\left(\frac{A}{s}\right)^j\left(1-\frac{A}{s}\right)^{n-j}\epsilon_{jk}\leq B.$$

It follows directly from (16) and (17) that the limit drifts are given by

$$c(\psi)=\lambda-t(A\psi)$$

$$e(\psi)=\beta+(\alpha-\beta)e^{-A\psi},$$

respectively, for $\psi \in [0, +\infty)$. Uniform convergence to the limit drifts follows immediately from the results given for the perfect state information case (Property 4). Also it is clear the $t(x)$ is differentiable (see (6), where $0 \le C_n \le n$). Therefore, properties ii) and iii) in Mikhailov's Theorem are satisfied.

In order to check property iv) note that if $\psi_0 = \xi$, then it follows from C1′ that

$$c(\psi_0) = \psi_0 e(\psi_0).$$

But, at that point, $c(\psi_0) < 0$ because of the choice of ξ. There is no other root of the equation $c(\psi) = \psi e(\psi)$, and, therefore, property v) follows. To see this, note that because of C1′, $c(\psi) = \psi e(\psi)$ for $\psi \ne \xi$ is equivalent to

$$\beta = \frac{\frac{\lambda - t(A\psi)}{\psi} - e^{A(\xi-\psi)}\frac{\lambda - t(A\xi)}{\xi}}{1 - e^{A(\xi-\psi)}}$$

which is impossible if $\psi \ne \xi$ because of C2′. □

It can be shown [9] that $m_\xi(\lambda)$ is finite for all nonnegative λ and ξ, and therefore the set of control laws defined by C1′ and C2′ is nonempty. Actually, the set of control laws in Theorem 4 is a subset of those in Theorem 5 because in Theorem 5 we can choose $\xi = 1$, in which case C2 is equivalent to C1′ and C1 is more restrictive than C2′ because $\lambda \ge m_1(\lambda)$ [9]. □

V. Conclusion

In this paper we have investigated the properties of decentralized control algorithms for a random access channel with multipacket reception capability. By using the working hypothesis that the users are aware of the value of the backlog, we have determined the best throughput achievable by any such protocol, as well as a simple way to achieve it. The optimum throughput has been shown to be given by the maximum average number of successes per slot when the number of transmissions, per slot is Poisson distributed. In the imperfect state information case, we have shown that the same throughput achieved in the perfect state information case can be achieved by using in lieu of the true backlog, an estimate of the backlog computed at each station using binary feedback, and we have used this estimate to derive a control scheme which is optimal in the sense that it achieves the optimal throughput determined earlier. This is true provided the reception matrix verifies condition C0, which puts some restrictions on the number of successes per slot. By using Mikhailov's result, C0 can be replaced by the weaker condition C0′. In this case however, geometric ergodicity was not ensured. Note that the feedback empty/nonempty used in Sections III and IV may be less than the available feedback in many practical situations, but no further information is needed: a ternary feedback would not shorten the proof or achieve better throughput.

Finally, let us mention that one can easily modify the proof of Theorem 4 to show that a similar result holds with the IFT access rule. More precisely, under a hypothesis paralleling those of Theorem 4, one can build a control scheme based on a binary feedback empty/nonempty such that the Markov vector (X_t, S_t) is geometrically ergodic for $\lambda < \sup_{x \ge 0} T(x)$. Using Theorem 3, it can be seen that the maximum stable throughput is the same for both access rules when the new packet arrivals are Poisson distributed.

Appendix

Kaplan's Condition

Consider a Markov chain with denumerable state-space D, and one-step transition probability matrix $(P_{xy})_{(x,y)\in D}$. Let $V(x)$ by a Lyapunov function on D. Then the generalized Kaplan's condition holds if there exists a positive constant B such that for all $z \in [0, 1[$ and all $x \in D$

$$z^{V(x)} - \sum_{y\in D} P_{xy} z^{V(y)} \ge -B(1-z).$$

1) One-Dimensional Kaplan's Condition: Consider the model of Section II with a control scheme $p_n = F(X_n)$, and the Lyapunov function $V(x) = x$. To check Kaplan's condition, it is enough from [27] to show that the downward part of the drift $-D(i) = \Sigma_{k=1}^{i} kP_{i,i-k}$ is bounded below. For $i \ge 1$ and $1 \le k \le i$ we have

$$P_{i,i-k} = \sum_{n=0}^{i-k} \lambda_n \sum_{j=k+n}^{i} \binom{i}{j} F(i)^j (1-F(i))^{i-j} \epsilon_{j,k+n}.$$

After a change of variable, it follows that

$$D(i) = \sum_{j=1}^{i} \binom{i}{j} F(i)^j (1-F(i))^{i-j} \sum_{n=0}^{j-1} \lambda_n \sum_{k=n+1}^{j} (k-n)\epsilon_{j,k}. \tag{A-1}$$

If $(C_n)_{n\ge1}$ is bounded, then Kaplan's condition holds independent of the retransmission policy. Denoting by B_c an upper bound for $(C_n)_{n\ge1}$, (A-1) becomes

$$-D(i) \ge -\sum_{j=1}^{i} \binom{i}{j} F(i)^j (1-F(i))^{i-j} \sum_{n=0}^{j-1} \lambda_n C_j$$

$$\ge -\sum_{j=1}^{i} \binom{i}{j} F(i)^j (1-F(i))^{i-j} C_j \ge -B_c. \tag{A-2}$$

2) Two-Dimensional Kaplan's Condition: Consider now the multipacket channel with a general control algorithm (1). Then (X_t, S_t) is the Markov chain of interest, and the relevant Lyapunov function is $V(n, s) = n$. We prove again that Kaplan's condition holds provided that $(C_n)_{n\ge1}$ is bounded. From [27], it is enough also in this case to show that the downward part $T(x)$ of the generalized drift is bounded below, with $T(x) = \Sigma_{y/V(y)<V(x)} P_{xy} (V(y) - V(x))$. Given a state $x = (i, s)$, we have

$$T(x) = -\sum_{r=1}^{i} r \sum_{k} P[X_{n+1} = i-r,\ S_{n+1} = k \mid X_n = i,\ S_n = s]$$

$$= -\sum_{r=1}^{i} rP[X_{n+1} = i-r \mid x_n = i,\ S_n = s]$$

which is, in the same way as before

$$T(x) = -\sum_{r=1}^{i-r} r \sum_{n=0}^{i-r} \lambda_n \sum_{j=r+n}^{i} \binom{i}{j} (F(s))^j (1-F(s))^{i-j} \epsilon_{j,r+n}$$

$$= -\sum_{j=1}^{i} \binom{i}{j} F(s)^j (1-F(s))^{i-j} \sum_{n=0}^{j-1} \lambda_n \sum_{r=n+1}^{j} (r-n)\epsilon_{jr}$$

this expression is similar to (A-1), and the end of the proof is the same as in (A-2).

References

[1] N. Abramson, "The throughput of packet broadcasting channels," *IEEE Trans. Commun.*, vol. COM-25, pp. 117-128, 1977.

[2] J. C. Arnbak and W. van Blitterswijk, "Capacity of slotted Aloha in Rayleigh fading channels," *IEEE J. Select. Areas Commun.*, vol. SAC-5, no. 2, pp. 261-269, 1987.

[3] D. H. Davis and S. A. Gronemeyer, "Performance of slotted Aloha random access with delay capture and randomized time of arrival," *IEEE Trans. Commun.*, vol. COM-28, pp. 703-710, 1980.

[4] G. Fayolle, E. Gelenbe, and J. Labetoulle, "Stability and optimal control of the packet switching broadcast channel," *J. Ass. Comput. Mach.*, vol. 24, no. 3, pp. 375-386, 1977.

[5] F. G. Foster, "On the stochastic matrices associated with certain queuing processes," *Ann. Math. Stat.*, no. 24, pp. 355-360, 1953.
[6] S. Ghez, S. Verdú, and S. C. Schwartz, "Stability of multipacket Aloha," in *Proc. 21st Conf. ISS*, The Johns Hopkins Univ., Baltimore, MD, Mar. 1987.
[7] ——, "On decentralized control algorithms for multipacket Aloha," in *Proc. 25th Allerton Conf. Commun., Contr., Comput.*, Oct. 1987.
[8] ——, "Stability properties of slotted Aloha with multipacket reception capability," *IEEE Trans. Automat. Contr.*, vol. 33, July 1988.
[9] S. Ghez, "Random access communications for the multipacket channel," Ph.D dissertation, Dept. Elec. Eng., Princeton Univ., 1989.
[10] D J Goodman and A A M. Saleh, "The near/far effect in local Aloha radio communications," *IEEE Trans. Vehic. Technol.*, vol. VT-36, no. 1, Feb. 1987.
[11] B. Hajek, "Hitting-time and occupation-time bounds implied by drift analysis with applications," *Adv. Appl. Prob.*, vol. 14, pp. 502-525, 1982.
[12] B. Hajek and T. Van Loon, "Decentralized dynamic control of a multiaccess broadcast channel," *IEEE Trans. Automat. Contr.*, vol. AC-27, pp. 559-569, 1982.
[13] M Kaplan, "A sufficient condition for nonergodicty of a Markov chain," *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 470-471, 1979.
[14] J. G. Kemeny, J. L. Snell, and A. W. Knapp, *Denumerable Markov Chains*. New York: Springer-Verlag, 1976.
[15] F. P. Kelly, "Stochastic models of computer communication systems," *J. R. Statist. Soc. B*, vol. 47, no. 3, pp. 379-395, 1985.
[16] C. C. Lee, "Random signal levels for channel access in packet ra networks," *IEEE J. Select. Areas Commun.*, vol. SAC-5, no. 6, pp. 1026-1034, 1987.
[17] N. Mehravari, "Collision resolution in random access systems with multiple reception," preprint, 1987.
[18] J. J. Metzner, "On improving utilization in Aloha networks," *IEEE Trans. Commun.*, vol. COM-24, pp. 447-448, Apr. 1976.
[19] V. A. Mikhailov, "Geometrical analysis of the stability of Markov chains in R_n^+ and its application to throughput evaluation of the adaptive random access algorithm," *Problemy Peredachi Informatsii*, vol. 24, no. 1, pp. 61-73, Jan.-Mar. 1988; translated in *Problems of Inform. Transmission*, July 1988.
[20] C. Namislo, "Analysis of mobile radio slotted Aloha networks," *IEEE J.Select. Areas Commun.*, vol. SAC-2, pp. 583-588, 1984.
[21] A. G. Pakes, "Some conditions for ergodicity and recurrence of Markov chains," *Operat. Res.*, vol. 17, pp. 1058-1061, 1969.
[22] A Polydoros and J, Silvester, "Slotted random access spread-spectrum networks: An analytical framework," *IEEE J. Select. Areas Commun.*, vol. SAC-5, no. 6, pp. 989-1002, 1987.
[23] B. Ramamurthi, A. A. M. Saleh, and D. J. Goodman, "Aloha with perfect capture," preprint 1986.
[24] D. Raychaudhuri, "Performance analysis of random access packet-switched code division multiple access systems," *IEEE Trans. Commun.*, vol. COM-29, pp. 895-901, 1981.
[25] R. L. Rivest, "Network control by Bayesan broadcast, ' *IEEE Trans. Inform. Theory*, vol. IT-33, no. 3, pp. 323-328, 1987.
[26] L G Roberts, "Aloha packet system with and withpout slots and capture," *Comput. Commun. Rev.*, no. 5, pp. 28-42, 1975.
[27] L I Sennott, "Tests for the nonergodicity of multidimensional Markov chains," *Operat. Res.*, vol. 33, pp. 161-167, 1985.
[28] A. Schwartz and M Sidi, "Erasure capture and noise errors in controlled multiple access networks." in *Proc. 25th Conf. Decision Conf.*, Dec. 1986, pp. 1333-1334.
[29] M. K. Simon, J. K. Omura, R. A. Scholtz, and B. K. Levitt, *Spread-Spectrum Communications*. New York. Computer Science Press, 1985.
[30] J N Tsitsiklis, "Analysis of a multiaccess control scheme," *IEEE Trans. Automat. Contr.*, vol. AC-32, pp. 1017-1020, 1987.
[31] B. S. Tsybakov, V. A. Mikhailov, and N. B. Likhanov, "Bounds for packet transmission rate in a random multiple access system," *Problemy Peredachi Informatsii*, vol. 19, no. 1, pp. 61-81, 1983.
[32] R. L. Tweedie, "Sufficient conditions for ergodicity and recurrence of Markov chains on a general state space," *Stoch. Proc. Appl.*, no. 3, pp. 385-403, 1975.
[33] S. Verdú, "Minimum probability of error for asynchronous Gaussian multiple-access channels," *IEEE Trans. Inform. Theory*, vol. IT-32, pp. 85-96, 1986.
[34] J. E. Wieselthier, A. Ephremides, and L. A. Michaels, "An exact analysis and performance evaluation of framed Aloha with capture," preprint, 1987.
[35] A. Ephremides and S. Verdú, "Control and optimization methods in communication network problems," *IEEE Trans. Automat. Contr.*, vol. 34, pp. 930-942, Sept. 1989.

Sylvie Ghez (S'80) was born in Paris, France, on June 3, 1961. She received the engineering diploma in 1984 from the Ecole Nationale Supérieure des Télécommunications, Paris, and the M.S. degree in electrical engineering from Princeton University, Princeton, NJ, in 1987.

At present she is a doctoral candidate at Princeton University. Her research interests include multiple access protocols, network routing problems, queueing theory, and applied probability theory.

Ms. Ghez was the recipient of the Princeton Wallace Memorial Fellowship for the year 1987-1988.

Sergio Verdú (S'80-M'84-SM'88), for a photograph and biography, see p. 942 of the September 1989 issue of this TRANSACTIONS.

Stuart C. Schwartz (S'64-M'66-SM'83) was born in Brooklyn, NY, on July 12, 1939. He received the B.S. and M.S. degrees in aeronautical engineering from the Massachusetts Institute of Technology, Cambridge, in 1961, and the Ph.D. degree from the Information and Control Engineering Program, University of Michigan, Ann Arbor, in 1966.

While at M.I.T. he was associated with the Naval Supersonic Laboratory and the Instrumentation Laboratory. During the year 1961-1962 he was at the Jet Propulsion Laboratory, Pasadena, CA, working on problems in orbit estimation and telemetry. He is presently a Professor of Electrical Engineering and Chairman of the department at Princeton University, Princeton, NJ. He served as Associate Dean of the School of Engineering during the period July 1977-June 1980. During the academic year 1972-1973 he was a John S. Guggenheim Fellow and a Visiting Associate Professor in the Department of Electrical Engineering, Technion—Israel Institute of Technology, Haifa, Israel. During the academic year 1980-1981, he was a member of the Technical Staff at the Radio Research Laboratory, Bell Telephone Laboratories, Crawford Hill, NJ. His principal research interests are in the application of probability and stochastic processes to problems in statistical communication and system theory.

Dr. Schwartz is a member of Sigma Gamma Tau, Eta Kappa Nu, and Sigma Xi He has served as an Editor for *SIAM Journal on Applied Mathematics* and as Program Chairman for the 1986 ISIT.

Single-User Detectors for Multiuser Channels

H. Vincent Poor
Sergio Verdu

Reprinted from
IEEE TRANSACTIONS ON COMMUNICATIONS
Vol. 36, No. 1, January 1988

Single-User Detectors for Multiuser Channels

H. VINCENT POOR, FELLOW, IEEE, AND SERGIO VERDU, MEMBER, IEEE

Abstract—Optimum decentralized demodulation for asynchronous Gaussian multiple-access channels is considered. It is assumed that the receiver is the destination of the information transmitted by only one active user, and single-user detectors that take into account the existence of the other active users in the channel are obtained. This approach is in contrast to both conventional demodulation, which is fully decentralized but neglects the presence of multiple-access interference, and globally optimum demodulation, which requires centralized sequence detection. The problem considered is one of signal detection in additive colored non-Gaussian noise, and attention is focused on one-shot structures where detection of each symbol is based only on the received process during its corresponding interval. Particular emphasis is placed on asymptotically optimum detectors for each of the following situations: 1) weak interferers, 2) CDMA signature waveforms with long spreading codes, and 3) low background Gaussian noise level.

I. INTRODUCTION

THE conventional approach to the demodulation of code-division multiplexed multiuser digital communications is to demodulate each user as if it were the only user in the channel. The multiple-access capability of such systems is thus achieved by using complex signal constellations that exhibit favorable cross-correlation properties. (See, for example, [1] for a description of conventional multiple-access demodulation techniques.) However, recent work by Verdu [2] has shown that substantial performance gains can be achieved in coherent multiuser systems by using a receiver that takes advantage of the structure of the multiple-access interference. For example, this approach can be used to alleviate such limitations as the near/far problem in the direct-sequence spread-spectrum multiple-access (DS/SSMA) format. The performance gains realized by the receiver proposed in [2] are achieved by the use of simultaneous sequence detection of all users in the channel, a task that requires a centralized implementation and a high degree of software complexity (for example, the decision algorithm required is a dynamic program (DP) whose complexity is $O(2^K)$ where K is the number of users in the channel). Since the implementation costs of such fully centralized detection algorithms may be unacceptably high for many applications, and since network security restrictions may not permit the distribution of all user's signaling waveforms to all demodulating terminals, it is of interest to consider demodulators that lie between these two philosophies of conventional demodulation, in which other users' signaling waveforms to all demodulating terminals, it is optimum demodulation, in which all users in the channel are tracked and demodulated simultaneously. The performance results obtained in [2]–[4] indicate that an attractive compromise in practice is to use optimum multiuser demodulators for only a subset of the active users and simply neglect the presence of all other users. In order to take advantage of the superior performance achievable by multiuser detectors, the subset of active users to be taken into account at the receiver should contain all the users whose power is sufficient regardless of whether their messages are destined to that particular location.

In this paper we consider the case of full decentralization where the receiver is constrained to track and lock to the signal of only one user, but unlike the conventional single-user detector, is optimized to take into account the structure of multiple-access interference in making decisions. We consider several design approaches that can be used to optimize these structures depending on the amount of information that one is able to assume to be known about the signature waveforms assigned to the interfering users.

This paper is organized as follows. In Section II, we will first discuss the general structure of optimum decentralized demodulators that simultaneously track and demodulate a group of D users from a total population of K users sharing a common communications channel where $D \leq K$. We then consider the structure of optimum *single-user* detectors ($D = 1$), and particularize to the case $K = 2$ to illustrate this structure. The results of Section II are for general antipodal signaling formats In Section III, we turn to the development of specific results for single-user demodulation of DS/SSMA transmissions. In this modulation technique, which among coherent signaling formats is of particular practical interest, each signature waveform consists of a sequence of chip waveforms whose polarities are determined by a binary word assigned to each user. The specific structure of the direct-sequence format allows for the development of useful approximations to optimum single-user detection which are asymptotically exact as either the length of the spreading codes or the signal-to-background-noise ratio (SBNR) increases without bound. We also show that (with $K = 2$) even in the absence of any prior knowledge about either the spreading codes or the timing of the multiuser interference, the optimum single-user detector is *not* multiple-access noise-limited as the background thermal noise level vanishes. This is in contrast to the conventional detector, which can incur an irreducible error probability even in the absence of background noise. All of these results for DS/SSMA require only that the chip waveform (which is usually common to all users in a given network) be known. Thus, these techniques can be applied in secure networks where the distribution of one user's spreading code to other users is not desirable.

In Section IV, we return to general coherent signaling formats to consider the problem of optimum single-user detection in the presence of weak unlocked interfering users. We model this problem by assuming that the multiple-access interference is multiplied by a small amplitude factor ϵ. We then derive an expression for the likelihood ratio statistic for optimum symbol decisions on the locked user that is of the form of the conventional correlation statistic, modified by an ϵ^2 term involving signal cross-correlation functions, and then having higher order terms of order ϵ^4. The resulting locally optimum detector correlates the observation with a replica of the waveform of the user of interest, suitably smoothed to take into account the presence of multiple-access interference.

Paper approved by the Editor for Spread Spectrum of the IEEE Communications Society. Manuscript received February 17, 1987; revised July 27, 1987 This work was supported in part by the U.S. Army Research Office under Contract DAAL03-87-K-0062.

H V Poor is with the Department of Electrical and Computer Engineering and the Coordinated Science Laboratory, University of Illinois, Urbana, IL 61801.

S. Verdu is with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544.

IEEE Log Number 8717473.

0090-6778/88/0100-0050$01.00

II. Optimum Decentralized Detection for Multiuser Channels

Throughout this paper, we consider a received signal model of the form

$$r_t = S_t(b) + n_t, \qquad -\infty < t < \infty \tag{2.1}$$

where $\{n_t;\ -\infty < t < \infty\}$ represents white Gaussian noise with spectral height $N_0/2$, and where the received signal $S_t(b)$ is the superposition of transmissions received from K separate asynchronous users, i.e.,

$$S_t(b) = \sum_{k=1}^{K} \sum_{i=-M}^{M} b_k(i) s_k(t - iT - \tau_k) \tag{2.2}$$

where T is the symbol interval, $b_k(i)$ is the ith symbol of the kth user, τ_k is the relative delay (modulo T) with which the kth user's transmission is received, and $s_k(t)$ is the signature waveform assigned to the kth user. (It is assumed that $s_k(t)$ is zero for $t \notin [0, T]$.) Note that $(2M + 1)$ is the number of symbols per user in the given transmission, and b denotes the $K \times (2M + 1)$ matrix whose (k, i) entry is $b_k(i)$.

Suppose that we wish to demodulate some group of D users from the total population of K users where $D \leq K$. For simplicity of notation, we assume that these D users of interest are labeled 1–D. Thus, we know $s_1, \cdots, s_D$ and $\tau_1, \cdots, \tau_D$, and the *maximum likelihood* demodulator chooses a symbol matrix $b_D \equiv \{b_k(i);\ k = 1, \cdots, D\}_{i=-M}^{M}$ to maximize the log-likelihood function

$$\frac{2}{N_0}\int_{-\infty}^{\infty} r_t S_t^D(b_D)\, dt - \frac{1}{N_0}\int_{-\infty}^{\infty} [S_t^D(b_D)]^2\, dt + \log E\left\{\exp\left[\frac{2}{N_0}\int_{-\infty}^{\infty} [r_t - S_t^D(b_D)] S_t^{MA}\, dt - \frac{1}{N_0}\int_{-\infty}^{\infty} [S_t^{MA}]^2\, dt\right]\right\} \tag{2.3}$$

where

$$S_t^D(b_D) = \sum_{k=1}^{D} \sum_{k=-M}^{M} b_k(i) s_k(t - iT - \tau_k), \qquad -\infty < t < \infty \tag{2.4}$$

$$S_t^{MA} = S_t(b) - S_t^D(b_D) \tag{2.5}$$

and where the expectation is over the ensemble of all unknown quantities in S_t^{MA}, including delays, symbols, and (possibly) waveforms.

Note that, even if we ignore the complexity of computing the expectation in (2.3), the time-complexity-per-demodulated-bit (TCB) of brute-force maximization of the log-likelihood function is $O(|A|^{D(2M+1)}/DM|A|)$ where $|A|$ is the size of the symbol alphabet. Thus, unless some simpler algorithm can be found, simultaneous maximum likelihood sequence detection of D users is out of the question from a practical point of view. For the particular case of fully centralized detection ($K = D$), it turns out that a much simpler algorithm can indeed be found. In particular, for this case, the expectation term in (2.3) disappears (since $S_t^{MA} \equiv 0$), and the remaining terms can be decomposed in a way that allows maximization of (2.3) with a dynamic programming algorithm yielding a TCB of $O(|A|)^K$. (See [2], [3] for details of this analysis.)

Unfortunately, for $D < K$, the decomposition of (2.3) necessary for a dynamic programming solution is not possible because of the coupling among symbols in the expectation term. This means that maximum likelihood sequence detection is not generally computationally feasible (its TCB is exponential in the number of symbols per user) if all users' signaling waveforms are not known and locked. Thus, in considering decentralized demodulators in multiuser channels, we will restrict our attention to algorithms which demodulate only a single symbol at a time, i.e., we consider *one-shot detectors.* We also will restrict attention to the binary signaling case, in which $b_k(i) \in \{-1, +1\}$ for all i, k. Extensions to general alphabets are, in most cases, straightforward.

In the sequel, we will consider the case of full decentralization, i.e., single-user detection, which can be modeled by the binary hypothesis-testing problem

$$H_0:\ r_t = s_1(t) + S_t^{MA} + n_t, \qquad 0 \leq t \leq T$$

$$H_1:\ r_t = -s_1(t) + S_t^{MA} + n_t, \qquad 0 \leq t \leq T \tag{2.6}$$

where $\{n_t;\ -\infty < t < \infty\}$ is the white Gaussian noise and where

$$S_t^{MA} = \sum_{k=2}^{K} [b_k^L s_k(t - \tau_k + T) + b_k^R s_k(t - \tau_k)], \qquad 0 \leq t \leq T \tag{2.7}$$

with b_k^L and b_k^R denoting the kth user's bits in the intervals $[-T + \tau_k, \tau_k)$ and $[\tau_k, T + \tau_k)$, respectively. We also assume that the receiver is coherent with user 1 so that $\{s_1(t);\ t \in [0, T]\}$ is a deterministic waveform, and that each user's signaling waveform is of the form

$$s_k(t) = (2w_k)^{1/2} a_k(t) \cos(\omega_c t + \theta_k) \tag{2.8}$$

where ω_c is known. We assume that $(\omega_c T/2\pi)$ is an integer large enough so that integrals of $2\omega_c$ components can be neglected.

Optimum (maximum likelihood/minimum error probability) decisions for (2.6) are based on comparing the likelihood ratio to a threshold. With this in mind, we give the following result.

Proposition 2.1: Suppose that the phase vector of the interfering users $\theta = (\theta_2, \cdots, \theta_K)$ is uniformly distributed on $[0, 2\pi)^{K-1}$ and is independent of $b_k = (b_k^L, b_k^R)$, $k = 2, \cdots, K$, $\tau = (\tau_2, \cdots, \tau_K)$ and $(a_k(t),\ t \in [0, T],\ k = 2, \cdots, K)$. If the dependence of $\|S^{MA}\|$ on θ can be neglected,[1] then the likelihood ratio for (2.6) can be written in the following form

$$\exp\left[\frac{4w_1^{1/2}}{N_0}\int_0^T r_p(t) a_1(t)\, dt\right] \frac{E\left[\prod_{k=2}^{K} I_0((\rho_k^2(b_k, \tau_k, 1) + \psi_k^2(b_k, \tau_k))^{1/2}) \exp\left(-\sum_{j=2}^{k-1} \Gamma_{kj}(b_k, b_j, \tau_k, \tau_j)\right)\right]}{E\left[\prod_{k=2}^{K} I_0((\rho_k^2(b_k, \tau_k, -1) + \psi_k^2(b_k, \tau_k))^{1/2}) \exp\left(-\sum_{j=2}^{k-1} \Gamma_{kj}(b_k, b_j, \tau_k, \tau_j)\right)\right]} \tag{2.9}$$

where the expectation is over the ensemble of bits, delays, and possibly waveforms of the interfering users[2] and we use the notation

$$\rho_k(b_k, \tau_k, e) = \frac{2(w_k)^{1/2}}{N_0}\int_0^T \alpha_k(b_k^L, b_k^R, t - \tau_k) \cdot [r_p(t) - e w_1^{1/2} a_1(t)]\, dt \tag{2.10}$$

[1] For a waveform $x = \{x(t),\ 0 \leq t \leq T\}$, $\|x\|^2$ denotes $\int_0^T x^2(t)\, dt$

[2] $I_0(\ )$ is the modified Bessel function of the first kind of order 0, i.e., $I_0(x) = 1/2\pi \int_0^{2\pi} \exp(x \cos\theta)\, d\theta$.

$$\psi_k(b_k, \tau_k)=\frac{2(w_k)^{1/2}}{N_0}\int_0^T \alpha_k(b_k^L, b_k^R, t-\tau_k)r_q(t)\,dt \quad (2.11)$$

$$\Gamma_{kj}(b_k, b_j, \tau_k, \tau_j)=\frac{2(w_k w_j)^{1/2}}{N_0}\int_0^T \alpha_k(b_k^L, b_k^R, t-\tau_k)\cdot \alpha_j(b_j^L, b_j^R, t-\tau_j)\,dt \quad (2.12)$$

$$r_p(t)=\sqrt{2}r_t \cos(\omega_c t+\theta_1) \quad (2.13)$$

$$r_q(t)=\sqrt{2}r_t \sin(\omega_c t+\theta_1) \quad (2.14)$$

$$\alpha_k(b, c, \lambda)=ba_k(\lambda+T)+ca_k(\lambda). \quad (2.15)$$

Proof: The likelihood ratio for (2.6) is equal to the ratio of expected values of conditionally Gaussian *a priori* densities where the expected value is taken with respect to all random quantities in S^{MA}; this is given straightforwardly by

$$LR=\frac{\exp\left[-\frac{1}{N_0}\|r-s_1\|^2\right]}{\exp\left[-\frac{1}{N_0}\|r+s_1\|^2\right]}\cdot\frac{E\left\{\exp\left[-\frac{1}{N_0}\|S^{MA}\|^2+\frac{2}{N_0}\langle r-s_1, S^{MA}\rangle\right]\right\}}{E\left\{\exp\left[-\frac{1}{N_0}\|S^{MA}\|^2+\frac{2}{N_0}\langle r+s_1, S^{MA}\rangle\right]\right\}} \quad (2.16)$$

where, for functions x and y on $[0, T]$, the notation $\langle x, y\rangle$ denotes $\int_0^T x(t)y(t)\,dt$. The first ratio in the above expression is readily shown to be equal to $\exp\,[(4(w_1)^{1/2}/N_0)^T \int_0^T r_p(t)a_1(t)\,dt]$. Now, neglecting the dependence of $\|S^{MA}\|$ on θ, we have for every b and τ

$$\frac{1}{N_0}\|S^{MA}\|^2=\left[\sum_{k=2}^{K}\left(\frac{w_k}{N_0}+\sum_{j=2}^{k-1}\Gamma_{kj}(b_k, b_j, \tau_k, \tau_j)\right)\right]. \quad (2.17)$$

So, it remains to show that for all $(\alpha_k, b_k, \tau_k, k=2, \cdots, K)$, we have

$$\int_0^{2\pi}\cdots\int_0^{2\pi}\exp\left[\frac{2}{N_0}\langle r-ex_1, S^{MA}\rangle\right]\frac{d\theta_2\cdots d\theta_K}{(2\pi)^{K-1}}=\prod_{k=2}^{K} I_0\left((\rho_k^2(b_k, \tau_k, e)+\psi_k^2(b_k, \tau_k))^{1/2}\right). \quad (2.18)$$

To this end, we note that the following sequence of equalities holds:

$$\frac{2}{N_0}\langle r-es_1, S^{MA}\rangle$$
$$=\frac{2\sqrt{2}}{N_0}\sum_{k=2}^{K}\int_0^T (r(t)-es_1(t))w_k^{1/2}\alpha_k(b_k^L, b_k^R, t-\tau_k)\cdot\cos(\omega_c t-\omega_c\tau_k+\theta_k)\,dt$$
$$=\frac{2}{N_0}\sum_{k=2}^{K} w_k^{1/2}\int_0^T [(r_p(t)-2ew_1^{1/2}a_1(t)\cdot\cos^2(\omega_c t+\theta_1))\alpha_k(b_k^L, b_k^R, t-\tau_k)\cdot\cos(\theta_k-\omega_c\tau_k-\theta_1)-(r_q(t)-ew_1^{1/2}a_1(t)\cdot\sin(2\omega_c t+\theta_1))\alpha_k(b_k^L, b_k^R, t-\tau_k))\cdot\sin(\theta_k-\omega_c\tau_k-\theta_1)]\,dt$$
$$=\sum_{k=2}^{K}\rho_k(b_k, \tau_k, e)\cos(\theta_k-\omega_c\tau_k-\theta_1)-\psi_k(b_k, \tau_k)\sin(\theta_k-\omega_c\tau_k-\theta_1) \quad (2.19)$$

where the last equality follows by neglecting the integrals of the $2\omega_c$ terms. Equation (2.18) is immediate from (2.19) and the result follows. □

Note that the structure (2.9) consists of the single-user ($K=1$) correlation statistic

$$\frac{2w_1^{1/2}}{N_0}\int_0^T r_p(t)a_1(t)\,dt \quad (2.20)$$

used by conventional single-user receivers, modified by an additive correction term which accounts for the other users in the channel. Note that the received waveform enters this correction term through the sliding correlation statistics of (2.10) and (2.11).

The simplifying approximation in Proposition 2.1, which states that the energy of the multiple-access interference process is independent of the carrier phases, is certainly accurate when $\omega_c T$ is sufficiently large and the normalized (i.e, unit energy) cross correlations between the interfering users are low. We assume throughout this section and the following one that this independence is valid. If such an approximation is not assumed, then it can be shown straightforwardly that the multiplicative correction term in the likelihood ratio (2.16) is equal to

$$\frac{E\int_0^{2\pi}\cdots\int_0^{2\pi}\exp\left[\sum_{k=2}^{K}\rho_k(b_k, \tau_k, 1)\cos\alpha_k-\psi_k(b_k, \tau_k)\sin\alpha_k-\sum_{j=2}^{k-1}\Gamma_{kj}\cos(\alpha_k-\alpha_j)\right]d\alpha_2\cdots d\alpha_K}{E\int_0^{2\pi}\cdots\int_0^{2\pi}\exp\left[\sum_{k=2}^{K}\rho_k(b_k, \tau_k, -1)\cos\alpha_k-\psi_k(b_k, \tau_k)\sin\alpha_k-\sum_{j=2}^{k-1}\Gamma_{kj}\cos(\alpha_k-\alpha_j)\right]d\alpha_2\cdots d\alpha_K}. \quad (2.21)$$

Several variants of the general structure of (2.9) are of interest and will be considered here. One such variant is that in which the *modulation* waveforms of the interfering users $\{a_k(t);\ 2\le k\le K\}$ are known, and the remaining unknown quantities in $\{s_k(t);\ 2\le k\le K\}$ are all independent with the data bits and delays uniformly distributed in their ranges. In this case, the expectations in (2.9) reduce to

$$E\{(\cdot)\}=\frac{1}{(4T)^{K-1}}\sum_{b'\in 4^{K-1}}\cdot\int_{[0,T]^{K-1}}E\{(\cdot)|b', \tau\}\,d\tau_2\cdots d\tau_K \quad (2.22)$$

where $b'=(b_2, \cdots, b_K)$ and where the inner expectation is over the amplitudes. Thus, the computation of this likelihood ratio is of exponential complexity in K. Moreover, there will be a further substantial computational burden in computing the $(K-1)$-dimensional integral corresponding to averaging over the relative delays $\tau_2, \cdots, \tau_K$.

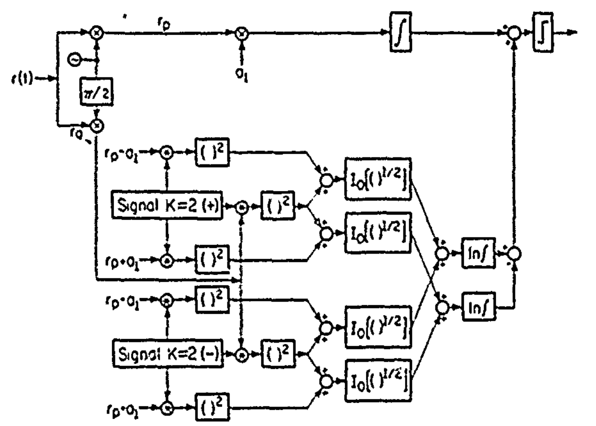


Fig. 1. Optimum single-user detector ($K = 2$). Replicas of $\alpha_2(+, +, -t)$ and $\alpha_2(+, -, -t)$ are generated by the blocks corresponding to the signal of the second user.

Fig. 1 illustrates the particularization of the demodulator derived in Proposition 2.1 to the two-user case. Note that the quadrature component of the input is used and that convolutions [required to generate (2.10) and (2.11)] and nonlinear memoryless operations are also needed.

For $K > 2$, the delay integrals do not appear to be obtainable in closed form. However, even if they could be, the exponential (in K) complexity of Σ_b, shows that optimum one-shot single-user detection in a K-user channel is at least as computationally burdensome as centralized simultaneous sequence detection of fully locked users. However, one-shot single-user detection does not require tracking phases, delays, and amplitudes of all users, and thus may be preferred if these quantities are not stable for relatively long periods of time. Moreover, and perhaps more importantly, Proposition 2.1 also applies to situations in which the modulating *waveforms* of the interferers $\{a_k(t); k = 2, \cdots, K\}$ are not known. This situation is the norm for the noncentral nodes in many practical radio networks, and thus the centralized detection algorithm of [2] cannot be applied to such cases unless the receiver estimates the unknown signal cross correlations. Furthermore, as it is shown in Section III, an important reduction in the complexity of computing (2.9) results from the modeling of the modulation waveforms of the interfering users as being signature sequences.

III. Single-User Detectors for DS/SSMA Channels

In practice, one of the most important types of code-division multiple-access systems is direct-sequence spread spectrum. This corresponds to the particular case of the model (2.1), (2.2), and (2.8), in which the kth user's signature waveform is of the form

$$a_k(t) = \sum_{i=0}^{N-1} c_{ki}\psi(t - iT_c), \qquad 0 \le t \le T \tag{3.1}$$

where $\{c_{ki}\}_{i=0}^{N-1}$ is a signature sequence of binary (± 1) digits, the chip waveform ψ is nonzero only on $[0, T_c]$, and the chip duration T_c is given by $T_c = T/N$. In many DS/SSMA multipoint-to-multipoint channels, it is frequently reasonable to assume that user 1 knows the chip waveforms of users 2–K, but not the specific signature sequences they employ. Since these sequences are usually chosen to be pseudonoise sequences, it is reasonable to model them (from the viewpoint of user 1) as independent sequences of independent, equiprobable binary digits. In this section, we apply this model for the interfering users in the likelihood ratio formula of Proposition 2.1. As we will see below, this affords a much more manageable form for the likelihood ratio in the limiting cases of practical interest, namely, when the number of chips is large and when the white Gaussian noise level is low.

A. Optimum Single-User Detection for Long Spreading Sequences

To study the large-N behavior of the likelihood ratio of Proposition 1, we first define the following functions:

$$\xi_e(\lambda) = \int_0^T [r_p(t) - e w_1^{1/2} a_1(t)]\psi(t - \lambda)\, dt \tag{3.2}$$

$$\phi(\lambda) = \int_0^T r_q(t)\psi(t - \lambda)\, dt \tag{3.3}$$

and

$$g_e(\lambda, \theta) = \xi_e(\lambda)\cos\theta - \phi(\lambda)\sin\theta \tag{3.4}$$

where the parameter e takes on the values $+1$ and -1 and g_e is abbreviated as g_+ and g_-, respectively, in the remainder of the section. Now fix $\tau \in (0, T)$ and suppose that $n \in \{1, \cdots, N\}$ is such that $(n - 1)T_c < \tau \le nT_c$. Then define

$$d_{kj} = \begin{cases} b_k^R & c_{kj-n} \qquad j - n \ge 0 \\ b_k^L & c_{kj-n+N} \qquad j - n < 0 \end{cases} \qquad j = 0, \cdots, N \tag{3.5}$$

and notice that $g_e(\tau - T + iT_c, \theta) = 0$ for $i \le N - n - 1$ and $g_e(\tau + iT_c, \theta) = 0$ for $i \ge N - n + 1$ because $\psi(t) = 0$ for $t \notin [0, T_c]$. Then it follows that

$$\begin{aligned}
&\frac{N_0}{2w_k^{1/2}}[\rho_k(b_k, \tau, e)\cos\theta - \psi_k(b_k, \tau)\sin\theta] \\
&= \sum_{i=0}^{N-1} c_{ki}[b_k^L g_e(\tau - T + iT_c, \theta) + b_k^R g_e(\tau + iT_c, \theta)] \\
&= c_{kN-n} b_k^L g_e(\tau - nT_c, \theta) + c_{kN-n} b_k^R g_e(\tau - nT_c + T, \theta) \\
&\quad + \sum_{i=1}^{N-1} d_{ki} g_e(\tau + (i - n)T_c, \theta) \\
&= \sum_{i=0}^{N} d_{ki} g_e(\tau + (i - n)T_c, \theta)
\end{aligned} \tag{3.6}$$

and thus, the distribution of $I_0((\rho_k^2(b_k, \tau_k, e) + \psi_k^2(b_k, \tau_k))^{1/2})$ is the same modulo T_c when τ_k is uniformly distributed.

Let us now consider the particular case of a single interferer $K = 2$ which may also be used to approximate the situation in which we have a single *dominant* interferer. In this case, the correction term of the likelihood ratio is equal to

$$\frac{\displaystyle\int_0^{T_c}\int_0^{2\pi} E\exp\left(\frac{2w_2^{1/2}}{N_0}\sum_{i=0}^{N} d_{2i} g_+(iT_c - \lambda, \theta)\right) d\theta\, d\lambda}{\displaystyle\int_0^{T_c}\int_0^{2\pi} E\exp\left(\frac{2w_2^{1/2}}{N_0}\sum_{i=0}^{N} d_{2i} g_-(iT_c - \lambda, \theta)\right) d\theta\, d\lambda} \tag{3.7}$$

where the expectation is over the independent and equiprobable sequence $d_{2i} \in \{-1, 1\}$, $i = 0, \cdots, N$. The integrands in the numerator and denominator of (3.7) are products of hyperbolic cosines which do not lend themselves to further simplification. However, if N (the number of chips) is large, the distribution of the discrete random variable $\Sigma_{i=0}^N d_{2i} g_e(iT_c - \lambda, \theta)$ approximates the normal curve, and further simplification of (3.7) is possible. To justify this approximation, we show that for each θ, λ and each realization of

$$x(t) = [r_p(t) - e w_1^{1/2} a_1(t)]\cos\theta - r_q(t)\sin\theta$$

such that $\sup\{|x(t)| \, t \in (0, T)\} < \infty$, the following triangular array of random variables[3]

$$\zeta_{ni} = d_{2i} \int_{(i-1)T/n}^{iT/n} x(t-\lambda)\, dt \qquad i = 1, \cdots, n \tag{3.8}$$

satisfies the Lindeberg–Feller condition (e.g., [5])

$$\lim_{n\to\infty} \sum_{i=1}^{n} E[(\zeta_{ni}/e_n)^2 I\{|\zeta_{ni}| > \delta e_n\}] = 0 \qquad \text{for every } \delta > 0 \tag{3.9}$$

where $e_n^2 = E[\Sigma_{i=1}^n \zeta_{ni}^2]$. To check (3.9), first note that

$$\lim_{n\to\infty} \frac{n}{T} \sum_{i=1}^{n} \zeta_{ni}^2 = \lim_{n\to\infty} \frac{n}{T} \sum_{i=1}^{n} x^2(t_i^n) T^2/n^2 = \|x\|^2 \tag{3.10}$$

where $t_i^n \in ((i - 1/n)T, (iT/n))$ and the first equation in (3.10) uses the mean-value theorem on the integral of (3.8). Therefore, for every $\epsilon > 0$, there exists n_0 such that for all $1 \le i \le n$ and $n > n_0$,

$$I\{|\zeta_{ni}| > \delta e_n\} < I\left\{\frac{n}{T}\zeta_{ni}^2 > \delta^2(\|x\|^2 - \epsilon)\right\}. \tag{3.11}$$

But for each $\mu > 0$, we can find n_1 such that for $n > n_1$, we have

$$I\left\{\frac{n}{T}|\zeta_{ni}|^2 > \mu\right\} \le I\left\{\frac{T}{n}\sup_t {}^2|x(t)| > \mu\right\} = 0. \tag{3.12}$$

Hence, only a finite number of terms on the left-hand side of (3.9) are nonzero, and since $\lim_{n\to\infty} \zeta_{ni}^2/e_n^2 = 0$, (3.9) follows.

If $\rho \cos\theta - \psi \sin\theta$ is a Gaussian random variable, then it is straightforward to check that

$$EI_0(\sqrt{\rho^2 + \psi^2}) = \exp\left(\frac{1}{4}(E[\rho^2] + E[\psi^2])\right) I_0\left(\sqrt{\frac{1}{16}(E[\rho^2] - E[\psi^2])^2 + \frac{1}{4}E^2[\rho\psi]}\right). \tag{3.13}$$

Hence, using the Gaussian approximation[4] to the distribution of $\Sigma_{i=0}^N d_{2i} g_e(iT_c - \lambda, \theta)$, the correction term in (3.7) reduces to

$$\frac{\int_0^{T_c} \exp\left(\frac{w_2}{N_0^2}(\Xi_+(\lambda) + \Phi(\lambda))\right) I_0\left(\frac{w_2}{N_0^2}\sqrt{(\Xi_+(\lambda) - \Phi(\lambda))^2 + \Theta_+^2(\lambda)}\right) d\lambda}{\int_0^{T_c} \exp\left(\frac{w_2}{N_0^2}(\Xi_-(\lambda) + \Phi(\lambda))\right) I_0\left(\frac{w_2}{N_0^2}\sqrt{(\Xi_-(\lambda) - \Phi(\lambda))^2 + \Theta_-^2(\lambda)}\right) d\lambda} \tag{3.14}$$

where

$$\Xi_e(\lambda) = \sum_{i=0}^{N} \xi_e^2(iT_c - \lambda) \tag{3.15}$$

$$\Phi(\lambda) = \sum_{i=0}^{N} \phi^2(iT_c - \lambda) \tag{3.16}$$

and

$$\Theta_e(\lambda) = 2\sum_{i=0}^{N} \xi_e(iT_c - \lambda)\phi(iT_c - \lambda). \tag{3.17}$$

This structure is illustrated in Fig. 2. Note that there is considerable simplification in this structure over that of Fig. 1. In particular, each of the "+" and "−" channels involves chip-matched filtering of the in-phase and quadrature components followed by chip-rate sampling, quadratic accumulation, memoryless nonlinear transformation, and integration over the offset λ. This latter operation can be implemented in parallel form by decimating an M/T_c-rate sampler (rather than a $1/T_c$-rate sampler) where M is the number of points taken in the numerical computation of the integral.

Further simplification of the correction term (2.9) is also possible in the case $K > 2$ by using the Gaussian approximation. We can obtain an expression similar to (3.14) where the integration is now over the hypercube $[0, T_c]^{K-1}$. Analogously to the case $K = 2$, for each $\theta = (\theta_2, \cdots, \theta_K)$ and $\tau = (\tau_2, \cdots, \tau_K)$, the distribution of

$$\sum_{k=2}^{K} \rho_k(b_k, \tau_k, e)\cos\theta_k - \psi_k(b_k, \tau_k)\sin \cdot \theta_k - \sum_{j=2}^{k-1} \Gamma_{kj}(b_k, b_j, \tau_k, \tau_j)$$

is approximately Gaussian, and since Γ_{kj} is uncorrelated with ρ_k, ψ_k, ρ_j, and ψ_j, both the numerator and the denominator of the correction term in (2.9) are approximated by

$$\int_{[0,T_c]}^{K-1} \cdots \int d\tau_2 \cdots d\tau_K E \cdot \exp\left(-\sum_{k=2}^{K}\sum_{j=2}^{k-1}\Gamma_{kj}(b_k, b_j, \tau_k, \tau_j)\right) \cdot \sum_{k=2}^{K} EI_0(\sqrt{\rho_k^2(b_k, \tau_k, e) - \psi_k^2(b_k, \tau_k)}).$$

Now, $E\exp(-\Sigma_{k=2}^K \Sigma_{j=2}^{k-1} \Gamma_{kj}(b_k, b_j, \tau_k, \tau_j))$ depends only on τ and on the chip waveform, and since if $X \sim N(0, \sigma)$, then $E\exp X = \exp \sigma^2/2$, we have

$$\begin{aligned}
E\exp\left(-\sum_{k=2}^{K}\sum_{j=2}^{k-1}\Gamma_{kj}(b_k, b_j, \tau_k, \tau_j)\right) &\approx \exp\left(\frac{1}{2}E\left[\left(\sum_{k=2}^{K}\sum_{j=2}^{k-1}\Gamma_{kj}(b_k, b_j, \tau_k, \tau_j)\right)^2\right]\right) \\
&= \exp\left(\frac{1}{2}\sum_{k=2}^{K}\sum_{j=2}^{k-1} E\Gamma_{kj}^2(b_k, b_j, \tau_k, \tau_j)\right) \\
&\approx \prod_{k=2}^{K}\prod_{j=2}^{k-1} \exp\left[\frac{w_k w_j}{N_0^2}\left(\frac{|\tau_k - \tau_j|^2 + (T_c - |\tau_k - \tau_j|)^2}{T_c}\right)\right]
\end{aligned} \tag{3.18}$$

[3] For the sake of notational simplicity, here we consider the case of a rectangular chip waveform. In this case, $\zeta_{Ni} = d_{2i} g_e(iT_c - \lambda, \theta)$.

[4] It should be noted that the use of a central-limit theorem here is quite different from the Gaussian approximations used in many previous analyses of conventional single-user receivers. Here, we do not claim that the multiple-access interference is asymptotically a white Gaussian process, however, we do show via the Lindeberg–Feller condition (3.9) that the decision statistics in (3.7) are conditionally Gaussian random variables as the number of chips per symbol goes to infinity.

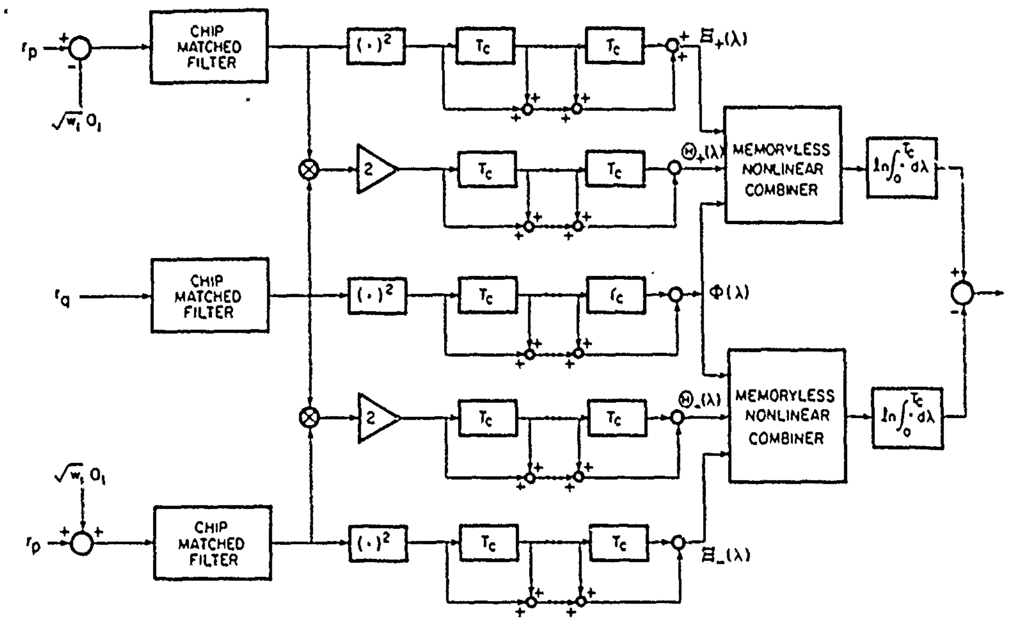


Fig. 2. Correction statistic for single-user detector with long spreading sequences.

where the last approximation follows by assuming that $\psi(t) = 1/T^{1/2}$ for $t \in [0, T_c]$ and by neglecting an $O(1/N^2)$ term in the exponent. Hence, the overall correction term is approximately equal to

$$\frac{\displaystyle\int_{[0,T_c]}\cdots\int_{[0,T_c]}\prod_{k=2}^{K}\exp\left(\frac{w_k}{N_0^2}(\Xi_+(\tau_k)+\Phi(\tau_k))\right)I_0\left(\frac{w_k}{N_0^2}\sqrt{(\Xi_+(\tau_k)-\Phi(\tau_k))^2+\Theta_+^2(\tau_k)}\right)\prod_{j=2}^{k-1}\exp\left[\frac{2w_kw_j}{NN_0^2}\left(\frac{|\tau_k-\tau_j|^2+(T_c-|\tau_k-\tau_j|)^2}{T_c}\right)\right]d\tau_2\cdots d\tau_K}{\displaystyle\int_{[0,T_c]}\cdots\int_{[0,T_c]}\prod_{k=2}^{K}\exp\left(\frac{w_k}{N_0^2}(\Xi_-(\tau_k)+\Phi(\tau_k))\right)I_0\left(\frac{w_k}{N_0^2}\sqrt{(\Xi_-(\tau_k)-\Phi(\tau_k))^2+\Theta_-^2(\tau_k)}\right)\prod_{j=2}^{k-1}\exp\left[\frac{2w_kw_j}{NN_0^2}\left(\frac{|\tau_k-\tau_j|^2+(T_c-|\tau_k-\tau_j|)^2}{T_c}\right)\right]d\tau_2\cdots d\tau_K}. \tag{3.19}$$

Notice that the term that couples the integrals in (3.19) is asymptotically independent of τ as $N \to \infty$. Hence, (3.19) approaches the product of $K-1$ (3.14)-like terms (substituting w_2 by w_k). Thus, in this limiting case, implementation of the multiuser correction term in the likelihood ratio involves the implementation of only one chip-matched-filter/quadratic-accumulator section followed by multiple averaging channels, one for each different value of w_k. Fig. 2 shows an implementation of the correction statistic to be added to the output of the single-user matched filter in the case of a single interferer. The general structure is the same, except that the memoryless nonlinearities output a process for each interferer which is then passed through a separate logarithmic integrator.

B. Optimum Single-User Detection for High SBNR

We now turn to another limiting case of the single-user detector for which a simplified form of the likelihood ratio exists, namely, the case when the power spectral density of the additive Gaussian noise goes to zero. In the above case, we saw that when the rest of the parameters are fixed, we can use a Gaussian approximation as $N \to \infty$. However, for fixed N, the error between the expected values of the exponentials, according to the true and Gaussian distributions, diverges as $N_0 \to 0$.[5] Hence, rather than using (3.19), we must take the limit as $N_0 \to 0$ of the original likelihood ratio (2.9). As in the previous analysis, we will first focus attention on the case of a single interferer ($K = 2$).

Since the spreading codes of the interfering users are modeled by the single-user receiver as equiprobable and independent binary sequences, the correction term of the likelihood ratio is given by (3.7) and the log-likelihood ratio is (except for a positive multiplicative constant) equal to

$$\begin{aligned}&2\int_0^T r_p(t)a_1(t)\,dt+\frac{N_0}{2w_1^{1/2}}\\&\cdot\log\int_0^{T_c}\int_0^{2\pi}E\exp\left(\frac{2w_2^{1/2}}{N_0}d_{2i}g_+(iT_c-\lambda,\theta)\right)d\theta\,d\lambda\\&-\frac{N_0}{2w_1^{1/2}}\log\int_0^{T_c}\int_0^{2\pi}\\&\cdot E\exp\left(\frac{2w_2^{1/2}}{N_0}\sum_{i=0}^{N}d_{2i}g_-(iT_c-\lambda,\theta)\right)d\theta\,d\lambda\end{aligned} \tag{3.20}$$

[5] This is due to the fact that as the variance goes to infinity, the error between the distributions accumulates on the tails (the true random variable is bounded) on which the expected value of the exponential largely depends.

where the expectation is over the independent and equiprobable sequence $d_{2i} \in \{-1, 1\}$, $i = 0, \cdots, N$. On taking the limit of the correction terms in (3.20), we obtain

$$\lim_{N_0 \to 0} \frac{N_0}{2w_1^{1/2}} \log \int_0^{T_c} \int_0^{2\pi} \cdot E \exp\left(\frac{2w_2^{1/2}}{N_0} \sum_{i=0}^{N} d_{2i} g_e(iT_c - \lambda, \theta)\right) d\theta\, d\lambda$$

$$= \lim_{N_0 \to 0} \frac{1}{w_1^{1/2}} \log \left[\int_0^{T_c} \int_0^{2\pi} \cdot \left[\exp\left(w_2^{1/2} \sum_{i=0}^{N} |g_e(iT_c - \lambda, \theta)|\right)\right]^{2/N_0} d\theta\, d\lambda \right]^{N_0/2}$$

$$- \lim_{N_0 \to 0} \frac{N_0}{2w_1^{1/2}} (N+1) \log 2$$

$$= \frac{1}{w_1^{1/2}} \log \sup_{\lambda, \theta} \left\{\exp\left(w_2^{1/2} \sum_{i=0}^{N} |g_e(iT_c - \lambda, \theta)|\right)\right\}$$

$$= (w_2/w_1)^{1/2} \sup_{\substack{\lambda \in [0, T_c] \\ \theta \in [0, 2\pi]}} \sum_{i=0}^{N} |g_e(iT_c - \lambda, \theta)|. \tag{3.21}$$

Therefore, in the limit as $N_0 \to 0$, the optimum single-user detector for $K = 2$ in the case of unknown interfering codes compares the test statistic

$$2 \int_0^T r_p(t) a_1(t)\, dt + (w_2/w_1)^{1/2} \sup_{\substack{\lambda \in [0, T_c] \\ \theta \in [0, 2\pi]}} \sum_{i=0}^{N} |g_+(iT_c - \lambda, \theta)|$$

$$- (w_2/w_1)^{1/2} \sup_{\substack{\lambda \in [0, T_c] \\ \theta \in [0, 2\pi]}} \sum_{i=0}^{N} |g_-(iT_c - \lambda, \theta)| \tag{3.22}$$

to a zero threshold. Note that as might be expected, (3.22) is also the limiting form of the generalized likelihood ratio test or maximum likelihood detector (see Helstrom [6, p. 291], for example).

We now investigate the error probability of the test in (3.22) when $N_0 = 0$. It was shown in [4] that when the delays, phases, and waveforms of all users are known, the fully centralized optimum detector achieves perfect demodulation with probability 1 in the absence of background noise. This is a nontrivial result, as is illustrated by the behavior of the *conventional* single-user detector which becomes multiple-access limited, i.e., the limit of its error probability as $N_0 \to 0$ is nonzero for sufficiently powerful interfering users. However, as in the present case, the conventional detector does not have access to the delays, phases, or signature sequences of the interfering users. So, the question arises as to whether an optimum single-user detector can achieve error-free performance regardless of the energies of the interfering users without knowledge of those parameters. The answer, in the two-user case, is given in the affirmative by the following result which does not put any restrictions on the signature sequences.

Proposition 3.1: Suppose $K = 2$ and $w_1 > 0$. If $r(t) = bs_1(t) + S^{MA}(t)$, $b \in \{-1, 1\}$, then

$$\operatorname{sgn}\left[2 \int_0^T r_p(t) a_1(t)\, dt + (w_2/w_1)^{1/2} \cdot \sup_{\substack{\lambda \in [0, T_c] \\ \theta \in [0, 2\pi]}} \sum_{i=0}^{N} |g_+(iT_c - \lambda, \theta)| - (w_2/w_1)^{1/2} \cdot \sup_{\substack{\lambda \in [0, T_c] \\ \theta \in [0, 2\pi]}} \sum_{i=0}^{N} |g_-(iT_c - \lambda, \theta)|\right] = b \tag{3.23}$$

with probability 1.

Proof: See Appendix.

In the general case of $K > 2$ users, the log-likelihood ratio is proportional to [cf. (2.16)]

$$2w_1^{1/2} \int_0^T r_p(t) a_1(t)\, dt + \frac{N_0}{2} \log \int_{[0, T_c]} \cdots \int_{K-1} \int_{[0, 2\pi]} \cdots \int_{K-1} \cdot E \exp\left[-\frac{1}{N_0} \|S^{MA}\|^2 + \frac{2}{N_0} \langle r - s_1, S^{MA} \rangle\right] d\theta\, d\tau$$

$$- \frac{N_0}{2} \log \int_{[0, T_c]}^{K-1} \cdots \int \int_{[0, 2\pi]}^{K-1} \cdots \int \cdot E \exp\left[-\frac{1}{N_0} \|S^{MA}\|^2 + \frac{2}{N_0} \langle r + s_1, S^{MA} \rangle\right] d\theta\, d\tau \tag{3.24}$$

where the expectation is over the independent sequences $d = \{d_{ki} \in \{-1, 1\};\ i = 0, \cdots, N,\ k = 2, \cdots, K\}$. As in (3.21), this expectation is dominated as $N_0 \to 0$ by the atom corresponding to the largest integrand, i.e.,

$$d^* \in \arg_d \max_d \Omega_e(d, \tau, \theta) \tag{3.25}$$

where

$$\Omega_e(d, \tau, \theta) = \langle r - es_1, S^{MA}(d) \rangle - \frac{1}{2} \|S^{MA}(d)\|^2 \tag{3.26}$$

and

$$S^{MA}(t, d) = \sum_{i=0}^{N} \sum_{k=2}^{K} d_{ki} (2w_k)^{1/2} \psi(t - (i-1)T_c - \tau_k) \cdot \cos(\omega_c t + \theta_k - \omega_c \tau_k). \tag{3.27}$$

Since there are $2^{K(N+1)}$ possible sequences, it is necessary to find an efficient way to carry out the maximization in (3.25). But (3.26) and (3.27) have the same structure as (2.3) and (2.4), respectively, so we can apply the results of [2] to carry out the maximization of (3.25) with linear complexity in N. On taking the limit of (3.24) as $N_0 \to 0$, we obtain the test statistic

$$2w_1^{1/2} \int_0^T r_p(t) a_1(t)\, dt + \sup_{\tau, \theta} \Omega_+^*(\tau, \theta) - \sup_{\tau, \theta} \Omega_-^*(\tau, \theta) \tag{3.28}$$

where $\Omega_e^*(\tau, \theta) = \Omega_e(d^*, \tau, \theta)$. Even if these quantities are obtained through efficient dynamic programming recursions as in [2], the main computational burden of (3.28) is the maximization over $[0, T_c]^{K-1}$ and $[0, 2\pi]^{K-1}$, which imposes severe limitations on its feasibility for even a moderate number of users. However, note that, in performing the maximization, the receiver is essentially acquiring the chip timing and carrier phases of the interfering users. Thus, in practice, it would normally be unnecessary to undergo a full search for the maximizing τ and θ in each symbol interval since these quantities will change little from symbol interval to symbol interval. For this reason, (3.28) might be reasonably efficient to implement in approximate form.

IV Locally Optimum Single-User Detectors with Weak Interferers

We have seen in the preceding sections that in multiple-access environments with many users, the complexity of optimum detection is increased considerably (over centralized reception) when the unwanted users are unlocked. This is true even without sequence detection and regardless of whether the interfering waveforms are known. However, one of the main incentives for the study of optimum decentralized detectors is the situation in which all or some of the interfering users are comparatively weak, so that it may be impractical to provide reliable synchronization for them. The objective of this section is to derive locally optimum (up to a third-order approximation) decentralized detectors for reception in the presence of weak unlocked users. As we shall see, such detectors can be viewed as versions of the detector that would be optimum without the weak interferers, modified to be robust against small deviations from the nominal white Gaussian noise statistics caused by weak multiple-access interference. As in the preceding sections, we consider both the case in which the waveforms of the interfering users are known, and the case in which they are coded with binary signature sequences unknown to the receiver. We will see here that the *locally optimum* version takes care only of the nonwhiteness of the multiple-access noise.

The approach we follow to derive locally optimum decentralized demodulators is to obtain an asymptotic form of the log-likelihood ratio for signal detection in contaminated white Gaussian noise given by the following result.

Lemma 4.1: Consider the following pair of statistical hypotheses:

$$\begin{aligned} H_0&: r_t = s_t^0 + \epsilon \tilde{n}_t + n_t \qquad t \in [t_p, t_f] \\ H_1&: r_t = s_t^1 + \epsilon \tilde{n}_t + n_t \qquad t \in [t_p, t_f] \end{aligned} \tag{4.1}$$

where s^1 and s^0 are deterministic finite-energy signals, $\{n_t\}$ is white Gaussian noise with spectral height σ^2, and $\{\tilde{n}_t, t \in [t_0, t_f]\}$ in a symmetric random process such that $\|\tilde{n}\| < B$ (a.s.) for some constant B, and whose correlation function is denoted by $C_{t,\lambda} = E[\tilde{n}_t \tilde{n}_\lambda]$, $(t, \lambda) \in [t_0, t_f]^2$. Then the log-likelihood ratio for (4.1) admits in the following expression:

$$\log LR(\epsilon) = \frac{1}{\sigma^2} \int_{t_0}^{t_f} \left[(s_t^1 - s_t^0) - \left(\frac{\epsilon}{\sigma}\right)^2 \int_{t_0}^{t_f} C_{t,\lambda}(s_\lambda^1 - s_\lambda^0)\, d\lambda \right] \left(r_t - \frac{1}{2} s_t^1 - \frac{1}{2} s_t^0 \right) dt + O(\epsilon^4). \tag{4.2}$$

Proof: Using the Cameron–Martin likelihood ratio formula, we obtain

$$\log LR(\epsilon) = \log \frac{D_1(\epsilon)}{D_0(\epsilon)} \tag{4.3}$$

where

$$D_i(\epsilon) = E\left[\exp\left(-\frac{1}{2\sigma^2} \left(\|s^i + \epsilon\tilde{n}\|^2 - 2\int_{t_p}^{t_f} (s_t^i + \epsilon \tilde{n}_t)\, dr_t \right) \right) \right] \qquad i = 0, 1 \tag{4.4}$$

where the expectation is over the ensemble of sample functions of $\{\tilde{n}_t, t \in [t_p, t_f]\}$.

In order to derive (4.2), we take the Taylor series expansion of (4.3) around the origin. Since $\tilde{n}_t$ is a symmetric random variable, it follows that $D_i(-\epsilon) = D_i(\epsilon)$, and hence the odd terms in the Taylor expansion of $D_i(\epsilon)|_{\epsilon=0}$ and $\log D_i(\epsilon)|_{\epsilon=0}$ are equal to zero.

Using the fact that $\|\tilde{n}\| < B$ a.s. and the Schwarz inequality, it follows that the expectation of every coefficient in the series expansion of the exponential in (4.4) exists, and we can write

$$D_i(\epsilon) = D_i(0) \left[1 + \frac{\epsilon^2}{2} E\left[\left(\frac{1}{\sigma^2} \int_{t_p}^{t_f} \tilde{n}_t (r_t - s_t^i)\, dt \right)^2 - \frac{1}{\sigma^2} \|\tilde{n}\|^2 \right] + O(\epsilon^4) \right]. \tag{4.5}$$

Now, since $\log(1 + x) = x + O(x^2)$, we obtain

$$\log \frac{D_1(\epsilon)}{D_0(\epsilon)} = \log \frac{D_1(0)}{D_0(0)} + \frac{\epsilon^2}{2} E\left[\left(\frac{1}{\sigma^2} \int_{t_p}^{t_f} \tilde{n}_t (r_t - s_t^1)\, dt \right)^2 - \left(\frac{1}{\sigma^2} \int_{t_p}^{t_f} \tilde{n}_t (r_t - s_t^0)\, dt \right)^2 \right] + O(\epsilon^4) \tag{4.6}$$

and (4.2) follows straightforwardly. □

Notice that the stringent condition $\|\tilde{n}\| < B$ (a.s.) allows a straightforward proof of Lemma 4.1 and is satisfied in the case in which we are interested, namely,

$$\tilde{n}_t = \sum_{i=-M}^{M} \sum_{k=D+1}^{K} b_k(i) s_k(t - iT - \tau_k);\ b_k(i) \in \{-1, 1\}. \tag{4.7}$$

If the waveforms $\{a_k(t), k = D + 1, \cdots, K\}$ are known by the receiver, then the autocorrelation function of $\tilde{n}$ with support in $\Re^2$ (for $M = \infty$) is equal to

$$C_{t,\lambda}^{MA} = \frac{1}{T} \cos(\omega_c(t - \lambda)) \sum_{k=D+1}^{K} w_k R_k(t - \lambda) \tag{4.8}$$

where

$$R_k(t) = \int_0^T a_k(s - t) a_k(s)\, ds. \tag{4.9}$$

If the waveforms of the interfering users have the form in (3.1) and the code of each user is unknown by the receiver and assumed to be equiprobably distributed among all $\{-1, 1\}$ sequences of length N, then the autocorrelation is

$$C_{t,\lambda}^{MA} = \frac{1}{T} \cos(\omega_c(t - \lambda)) \sum_{k=D+1}^{K} w_k \Psi(t - \lambda) \tag{4.10}$$

where the autocorrelation of the chip waveform is denoted by $\Psi(t) = \int_0^{T_c} \psi(s)\psi(s - t)\, dt$.

The one-shot single-user detector can be obtained readily from the result of Lemma 4.1. Since the signal of user 1 is antipodally modulated, we have

$$s_t^1 - s_t^0 = 2\sqrt{2w_1}\, a_1(t) \cos(\omega_c t + \theta_1)$$

and (4.2) becomes

$$\log LR = \frac{4\sqrt{w_1}}{N_0} \int_0^T r_p(t) \cdot \left[a_1(t) - \frac{1}{N_0 T} \sum_{k=2}^{K} w_k \int_0^T a_1(\lambda) R_k(t-\lambda)\, d\lambda \right] dt + O(\max_{k>1} w_k^2). \tag{4.11}$$

Hence, the locally optimum one-shot single-user detector is a conventional correlation receiver in which $a_1(t)$ is replaced by $a_1(t) - (1/N_0T) \Sigma_{k=2}^K w_k \int_0^T a_1(\lambda) R_k(t-\lambda)\, d\lambda$, $t \in [0, T]$, i.e., the pseudosignal is the output in $[T, 2T]$ of a causal linear filter, driven by $a_1(t)$, and whose impulse response is equal to $\delta(t-T) - (1/N_0T) \Sigma_{k=2}^K w_k R_k(t-T)$. If the signature sequences are unknown, the impulse response is $\delta(t - T) - (1/N_0T) \Sigma_{k=2}^K w_k \Psi(t - T)$, which amounts to a mild smoothing of the signal replica of the user of interest.

The locally optimum detector that locks to D of K users is, in fact, a generalization of this conclusion. Using Lemma 4.1, it can be shown (see [3, ch. 5] for details) that the locally optimum D-user detector is a centralized detector whose correlators use replicas of the unmodified waveforms of the users of interest. However, the input is processed by a causal filter that whitens the interference due to unlocked users, and whose impulse response depends on the autocorrelation function and signal-to-noise ratio of each interfering signal. This requires a modification of the DP algorithm to account for the intersymbol interference introduced by the prefilter, and results in a complexity of $O(2^{2D})$ as opposed to $O(2^D)$ for the corresponding algorithm that neglects the additional $K - D$ interferers.

V. Summary

In this paper, we have obtained decentralized single-user detectors which take into account the presence of interfering users. The general decentralized demodulation problem is one of sequence detection in additive colored non-Gaussian noise, and results in nonlinear detectors whose decision algorithms do not admit recursive forms and hence are more complex than their centralized counterparts. Important reductions in complexity occur when attention is focused on one-shot single-user detectors.

The general form of the single-user likelihood ratio obtained in Proposition 2.1 is equal to the single-user likelihood ratio affected by a correction term which depends on both the in-phase and quadrature components of the input. Both the case where the baseband interfering waveforms are known and the case where they are coded by an unknown signature sequence have been studied.

Under the assumption that the assigned waveforms are signature sequences with N chips per bit, we have obtained limiting forms of the correction term for $N \gg 1$ and for $N_0 = 0$. In the first case, the correction term depends on the received waveform only through the functions $\Xi_\pm(\lambda)$, $\Phi(\lambda)$, and $\Theta_\pm(\lambda)$ which represent the l_2 norms and inner product, respectively, of the subintegrals of an N partition (with offset $\lambda \in [0, T_c]$) of the in-phase and quadrature components of the received noise process under both hypotheses. The correction term when $N_0 = 0$ is best illustrated in the single-interferer case where it is obtained through the maximization over the relative phase and delay of the l_1 norm of the above subintegrals. It has been shown that this detector (which assumes knowledge of only the chip waveform and energy of the interfering user) achieves perfect demodulation in the absence of Gaussian noise regardless of the energy of the interference, thus avoiding the multiple-access limitation that plagues the conventional receiver. Using dynamic programming, the single-user detector can be implemented in linear time in N; however, its main computational burden is the maximization over $[0, T_c]^{K-1}$ and $[0, 2\pi]^{K-1}$ needed in the correction term.

Using an asymptotic form of the log-likelihood ratio for signal detection in contaminated white Gaussian noise, we have derived locally optimum detectors up to a third-order approximation in the amplitude of the interfering users. The locally optimum one-shot detector has been shown to be a single-user correlation receiver which uses a smooth replica of the signal of interest It has been shown in [3] that this approach can be generalized to the case of partial decentralization ($D > 1$), resulting in robustified versions of the centralized D-user receiver, which may offer substantial computational savings over the optimum K-user receiver.

Appendix

Proof of Proposition 3.1

We assume that the bit transmitted by user 1 is $b = 1$, the proof being identical in the antipodal case. For notational convenience and without lost of generality, we suppose that the relative delay of the interfering user is $0 < \tau_2 \leq T_c$; then it follows that

$$\alpha_2(b_2^L, b_2^R, t-\tau_2) = \sum_{i=0}^{N} d_i \psi(t - iT_c + \lambda_2) \tag{A.1}$$

where $\lambda_2 = T_c - \tau_2$, $d_0 = c_{2N-1} b_2^L$, and $d_{i+1} = c_{2i} b_2^R$ for $i = 0, N-1$. Let $\beta = \theta_1 + \omega_c \tau_2 - \theta_2$; then it is easy to show that

$$\int_0^T r_p(t) a_1(t) = w_1^{1/2} + w_2^{1/2} \cos\beta \cdot \int_0^T a_1(t) \alpha_2(b_2^L, b_2^R, t-\tau_2)\, dt \tag{A.2}$$

$$g_+(iT_c - \lambda, \theta) = w_2^{1/2} \cos(\theta + \beta) \cdot \int_0^T \alpha_2(b_2^L, b_2^R, t-\tau_2) \psi(t - iT_c + \lambda)\, dt \tag{A.3}$$

and

$$g_-(iT_c - \lambda, \theta) = 2 w_1^{1/2} \cos\theta \cdot \int_0^T a_1(t) \psi(t - iT_c + \lambda)\, dt + g_+(iT_c - \lambda, \theta). \tag{A.4}$$

We show now that

$$\sup_{\substack{\lambda \in [0, T_c] \\ \theta \in [0, 2\pi]}} \sum_{i=0}^{N} |g_+(iT_c - \lambda, \theta)| = w_2^{1/2}. \tag{A.5}$$

To that end, using (A.1) and (A.3), we obtain for every $\lambda \in [0, T_c]$

$$\sup_{\theta \in [0, 2\pi]} \sum_{i=0}^{N} |g_+(iT_c - \lambda, \theta)| = w_2^{1/2} \sum_{i=0}^{N} \left| \int_0^T \alpha_2(b_2^L, b_2^R, t-\tau_2) \psi(t - iT_c + \lambda)\, dt \right|$$

$$= w_2^{1/2} \sum_{i=0}^{N} \left| \int_0^T \sum_{j=0}^{N} d_j \psi(t - jT_c + \lambda_2) \psi(t - iT_c + \lambda)\, dt \right|$$

$$\leq w_2^{1/2} \int_0^T \sum_{i=0}^{N} \sum_{j=0}^{N} |\psi(t - jT_c + \lambda_2)| \left| \psi(t - iT_c + \lambda) \right| dt$$

$$\leq w_2^{1/2} \left(\int_0^T \sum_{j=0}^{N} \psi^2(t - jT_c + \lambda_2)\, dt \right)^{1/2} \cdot \left(\int_0^T \sum_{i=0}^{N} \psi^2(t - iT_c + \lambda)\, dt \right)^{1/2} = w_2^{1/2} \tag{A.6}$$

where the last two equations follow from the Schwarz inequality and from the relationship $\int_0^{T_c} (\psi^2(t + s) + \psi^2(t - T_c + s))\, dt = \int_0^{T_c} \psi^2(t) = 1/N,\ 0 \leq s \leq T_c$, respectively. But the right-hand side of (A.6) is achieved when $\lambda = \lambda_2$; hence, (A.5) follows. Consequently, in order to show that the sign of the log-likelihood ratio is positive, one has to prove that

$$2w_1 + w_2 + w_1^{1/2} w_2^{1/2} \int_0^T 2a_1(t) \alpha_2(b_2^L, b_2^R, t - \tau_2) \cdot \cos \beta\, dt - w_2^{1/2} \sum_{i=0}^{N} |g_-(iT_c - \lambda, \theta)| > 0 \tag{A.7}$$

for all $\lambda \in [0, T_c]$ and $\theta \in [0, 2\pi]$. Using (A.3) and (A.4), we obtain

$$2w_1 + w_2 + w_1^{1/2} w_2^{1/2} \int_0^T 2a_1(t) \alpha_2(b_2^L, b_2^R, t - \tau_2) \cos \beta\, dt - w_2^{1/2} \sum_{i=0}^{N} |g_-(iT_c - \lambda, \theta)|$$

$$= 2w_1 + w_2 + w_1^{1/2} w_2^{1/2} \cdot \int_0^T 2a_1(t) \alpha_2(b_2^L, b_2^R, t - \tau_2) \cos \beta\, dt - w_2^{1/2} \sum_{i=0}^{N} \left| \int_0^T (2w_1^{1/2} a_1(t) \cos \theta + w_2^{1/2} \cdot \alpha_2(b_2^L, b_2^R, t - \tau_2) \cos(\theta + \beta)) \psi(t - iT_c + \lambda)\, dt \right|. \tag{A.8}$$

The last term on the right-hand side of the above equation can be bounded as follows:

$$\sum_{i=0}^{N} \left| \int_0^T (2w_1^{1/2} a_1(t) \cos \theta + w_2^{1/2} \alpha_2(b_2^L, b_2^R, t - \tau_2) \cdot \cos(\theta + \beta)) \psi(t - iT_c + \lambda)\, dt \right|$$

$$\leq \sum_{i=0}^{N} \int_0^T |2w_1^{1/2} a_1(t) \cos \theta + w_2^{1/2} \alpha_2(b_2^L, b_2^R, t - \tau_2) \cdot \cos(\theta + \beta)| \cdot |\psi(t - iT_c + \lambda)|\, dt$$

$$= \int_0^T |2w_1^{1/2} a_1(t) \cos \theta + w_2^{1/2} \alpha_2(b_2^L, b_2^R, t - \tau_2) \cdot \cos(\theta + \beta)| \sum_{i=0}^{N} |\psi(t - iT_c + \lambda)|\, dt$$

$$\leq \left[\int_0^T (2w_1^{1/2} a_1(t) \cos \theta + w_2^{1/2} \alpha_2(b_2^L, b_2^R, t - \tau_2) \cdot \cos(\theta + \beta))^2\, dt \right]^{1/2} \left[\int_0^T \sum_{i=0}^{N} \psi^2(t - iT_c + \lambda)\, dt \right]^{1/2}$$

$$= \left[4w_1 \cos^2 \theta + w_2 \cos^2(\theta + \beta) + 4w_1^{1/2} w_2^{1/2} \cdot \cos \theta \cos(\theta + \beta) \int_0^T a_1(t) \alpha_2(b_2^L, b_2^R, t - \tau_2)\, dt \right]^{1/2}. \tag{A.9}$$

Since $\left| \int_0^T a_1(t) \alpha_2(b_2^L, b_2^R, t - \tau_2)\, dt \right| \leq 1$, we can denote $\int_0^T a_1(t) \alpha_2(b_2^L, b_2^R, t - \tau_2)\, dt = \cos \alpha$, and using (A.9), the right-hand side of (A.8) can be lower bounded by

$$2w_1 + w_2 + w_1^{1/2} w_2^{1/2} \int_0^T 2a_1(t) \alpha_2(b_2^L, b_2^R, t - \tau_2) \cos \beta\, dt - w_2^{1/2} \sum_{i=0}^{N} \left| \int_0^T (2w_1^{1/2} a_1(t) \cos \theta + w_2^{1/2} \cdot \alpha_2(b_2^L, b_2^R, t - \tau_2) \cos(\theta + \beta)) \psi(t - iT_c + \lambda)\, dt \right|$$

$$\geq 2w_1 + w_2 + 2w_1^{1/2} \cos \alpha \cos \beta - w_2^{1/2} (4w_1 \cos^2 \theta + w_2 \cos^2(\theta + \beta) + 4w_1^{1/2} w_2^{1/2} \cos \theta \cos(\theta + \beta) \cos \alpha)^{1/2}. \tag{A.10}$$

Now, since $2w_1 + w_2 + 2w_1^{1/2} w_2^{1/2} \cos \alpha \cos \beta > 0$, the sign of the right-hand side of (A.10) is equal to the sign of

$$(2w_1 + w_2 + 2w_1^{1/2} w_2^{1/2} \cos \alpha \cos \beta)^2 - [4w_1 w_2 \cos^2 \theta + w_2^2 \cos^2(\theta + \beta) + 4w_2 w_1^{1/2} w_2^{1/2} \cos \theta \cos \alpha \cos(\theta + \beta)]$$

$$= (2w_1 + 2w_1^{1/2} w_2^{1/2} \cos \alpha \cos \beta)^2 + 4w_1 w_2 (1 - \cos^2 \theta) + w_2^2 (1 - \cos^2(\theta + \beta)) + 4w_2 w_1^{1/2} w_2^{1/2} \cdot \cos \alpha [\cos \beta - \cos \theta \cos(\theta + \beta)]$$

$$= (2w_1 + 2w_1^{1/2} w_2^{1/2} \cos \alpha \cos \beta)^2 + 4w_1 w_2 \sin^2 \theta + w_2^2 \sin^2(\theta + \beta) + 4w_2 w_1^{1/2} w_2^{1/2} \sin \theta \sin(\theta + \beta) \cos \alpha$$

$$= (2w_1 + 2w_1^{1/2} w_2^{1/2} \cos \alpha \cos \beta)^2 + (w_2 \sin(\theta + \beta) \cos \alpha + 2\sqrt{w_1 w_2} \sin \theta)^2 + (w_2 \sin(\theta + \beta) \sin \alpha)^2. \tag{A.11}$$

Therefore, we have shown that (A.10) and, consequently, the left-hand side of (3.22), are nonnegative. Moreover, the right-hand side of (A.11) is equal to zero only if

$$2w_1 + 2w_1^{1/2} w_2^{1/2} \cos \alpha \cos \beta = 0, \tag{A.12}$$

but since $\beta = \theta_1 + \omega_c \tau_2 - \theta_2$ is uniformly distributed, (A.12) occurs with probability zero if $w_1 > 0$.

References

[1] M. K. Simon, J. K. Omura, R. A. Scholtz, and B. K. Levitt, *Spread Spectrum Communications, Vol. 3.* Rockville, MD: Computer Science Press, 1985.

[2] S. Verdu, "Minimum probability of error for asynchronous Gaussian multiple-access channels," *IEEE Trans. Inform. Theory,* vol. IT-32, pp. 85-96, Jan. 1986.
[3] ——, "Optimum multi-user signal detection," Ph.D. dissertation, Dep. Elec. Comput. Eng., Univ. Illinois, Urbana-Champaign, Coordinated Sci. Lab., Urbana, IL, Rep. T-151, Aug. 1984.
[4] ——, "Optimum multiuser asymptotic efficiency," *IEEE Trans. Commun.,* vol. COM-34, pp. 890-897, Sept. 1986.
[5] K. L. Chung, *A Course in Probability Theory,* 2nd ed. New York: Academic, 1974.
[6] C. W. Helstrom, *Statistical Theory of Signal Detection,* 2nd ed. London: Pergamon, 1968.

★

H. Vincent Poor (S'72-M'77-SM'82-F'87) was born in Columbus, GA, on October 2, 1951. He received the B.E.E. (with highest honor) and M.S. degrees in electrical engineering from Auburn University, Auburn, AL, in 1972 and 1974, respectively, and the M.A. and Ph.D. degrees in electrical engineering and computer science from Princeton University, Princeton, NJ, in 1976 and 1977, respectively.

Since 1977 he has been with the University of Illinois, Urbana-Champaign, where, since 1984, he has been Professor of Electrical and Computer Engineering and Research Professor in the Coordinated Science Laboratory. He was an Academic Visitor in the Department of Electrical Engineering at London University's Imperial College of Science and Technology in 1985, and a Visiting Professor in the Department of Electrical Engineering, University of Newcastle, NSW, Australia in 1987. His research interests are primarily in the general area of statistical signal processing in communications and control, with emphasis on the particular problems of robustness in signal detection, estimation, and filtering; processing techniques for multiple-access communications; quantization of inferential data; and modeling of nonstandard noise environments. He is the author of numerous publications in these areas including the book *An Introduction to Signal Detection and Estimation* (Springer-Verlag, 1987), a coeditor (with I. F. Blake) of the volume *Communications and Networks: A Survey of Recent Advances* (Springer-Verlag, 1986), and the Editor of the JAI Press series *Advances in Statistical Signal Processing.* He is also a member of the Editorial Board of the new Springer-Verlag journal, *Mathematics of Control, Systems and Signals.* He serves regularly as a consultant in the areas of digital communications, signal processing, and radar systems for several industrial and governmental clients.

Dr. Poor is a member of Phi Kappa Phi, Tau Beta Pi, Eta Kappa Nu, and Sigma Xi. He has served in several official capacities in various local, regional, and international IEEE organizations. In 1981-1982 he was Associate Editor for Estimation of the IEEE TRANSACTIONS ON AUTOMATIC CONTROL, and in 1982-1985 he was Associate Editor for Detection and Estimation of the IEEE TRANSACTIONS ON INFORMATION THEORY. He is currently serving as a member of the Board of Governors of the IEEE Information Theory Group, of which he was Treasurer in 1983-1985. Last year, he served as Program Chairman for the 25th IEEE Conference on Decision and Control held in Athens, Greece, in December 1986, and he will serve as General Chairman of the 1989 American Control Conference.

★

Sergio Verdu (S'80-M'84) was born in Barcelona, Catalonia, Spain, on August 15, 1958. He received the Telecommunications Engineer degree (first in class) from the Polytechnic University of Barcelona in 1980 and the M.S. and Ph.D. degrees in electrical engineering from the University of Illinois, Urbana-Champaign, in 1982 and 1984, respectively.

From 1981 to 1984 he was a Research Assistant in the Coordinated Science Laboratory of the University of Illinois. In September 1984 he joined the faculty of Princeton University, Princeton, NJ, where he is an Assistant Professor of Electrical Engineering. His current research interests are in the areas of multiuser communication and information theory and detection and estimation.

Dr. Verdu is a recipient of the National University Prize of Spain, of an IBM Faculty Development Award, and of the 1987 Rheinstein Outstanding Junior Faculty Award from the School of Engineering and Applied Science of Princeton University. He is a member of Sigma Xi and other professional societies.

T-COM/38/4//34846

Near—Far Resistance of Multiuser Detectors in Asynchronous Channels

Ruxandra Lupas
Sergio Verdú

Reprinted from
IEEE TRANSACTIONS ON COMMUNICATIONS
Vol. 38, No. 4, April 1990

Near-Far Resistance of Multiuser Detectors in Asynchronous Channels

RUXANDRA LUPAS, STUDENT MEMBER, IEEE, AND SERGIO VERDÚ, SENIOR MEMBER, IEEE

Abstract—We consider an asynchronous code-division multiple-access environment in which the receiver has knowledge of the signature waveforms of all the users. Under the assumption of white Gaussian background noise, we compare detectors by their worst case bit error rate in a low background noise near-far environment where the received energies of the users are unknown to the receiver and are not necessarily similar.

Conventional single-user detection in a multiuser channel is not near-far resistant, while the substantially higher performance of the optimum multiuser detector requires exponential complexity in the number of users. We explore suboptimal demodulation schemes which exhibit a low order of complexity while not exhibiting the impairment of the conventional single-user detector. Attention is focused on linear detectors, and it is shown that there exists a linear detector whose bit-error-rate is independent of the energy of the interfering users. Moreover it is shown that the near-far resistance of optimum multiuser detection can be achieved by a linear detector. The optimum linear detector for worst-case energies is found, along with existence conditions, which are always satisfied in the models of practical interest.

I. Introduction

THE near-far problem is the principal shortcoming of current radio networks using direct-sequence spread-spectrum multiple-access (DS/SSMA) communication systems. Those systems achieve multiple-access capability by assigning a distinct signature waveform to each user from a set of waveforms with low mutual crosscorrelations. Then, when the sum of the signals modulated by several asynchronous users is received, it is possible to recover the information transmitted by correlating the received process with replicas of the assigned signature waveforms. This demodulation scheme is conventionally used in practice, and its performance is satisfactory if two conditions are satisfied: first, the assigned signals need to have low crosscorrelations for all possible relative delays between the data streams transmitted by the asynchronous users, and second the powers of the received signals cannot be very dissimilar. If either of these conditions is not fulfilled, then the bit-error-rate and the antijamming capability of the conventional detector are degraded substantially. The reason why system performance is unacceptable when the received energies are dissimilar even with good (i.e., quasiorthogonal) signal constellations, is that the output of each correlator or matched filter contains a spurious component which is linear in the amplitude of each of the interfering users. Thus, as the multiuser interference grows, the bit-error-rate increases until the conventional detector is unable to recover the messages transmitted by the weak users.

Due to the severe reduction of the multiple-access capability and the increase of vulnerability to hostile sources caused by the near-far problem and its ubiquity in networks with dynamically changing topologies (such as mobile radio), its alleviation has been a target of researchers in the area for several years. However, success has been very limited and the only remedies implemented in practice have been to use power control or to design signals with more stringent crosscorrelation properties, which as we have noted, does not eliminate the near-far problem.

The viewpoint of this paper is that the near-far problem is not an inherent shortcoming of DS/SSMA systems, but of the conventional single-user detector. The optimum multiuser detector was obtained in [1] and was shown to be near-far resistant in the sense that a (very good) performance level can be guaranteed regardless of the relative energy of the transmitters. The optimum multiuser detector consists of a bank of matched filters and a Viterbi algorithm whose complexity is exponential in the number of users. In decentralized applications (where each receiver is only interested in demodulating the data sent by one transmitter), it is possible to drastically reduce the complexity of the optimum receiver (without compromising performance) by neglecting all but the comparatively powerful interferers. However, in this paper we propose a receiver (which we refer to as the decorrelating receiver) whose complexity is only linear in the number of users, and whose bit-error-rate is independent of the powers of the interferers at the receiver. Moreover, the decorrelating receiver achieves optimum near-far resistance (in a sense to be defined precisely in the sequel). The only requirement is the knowledge of the signature waveforms of the interfering users, and, in particular, no knowledge of the received energies is required, in contrast to the optimum receiver.

This paper generalizes the results obtained in [7] in the case of synchronous code-division multiple-access channels. Other recent attempts to derive detectors for multiuser channels include [9]–[11].

The multiple-access channel model considered in this paper is spelled out in Section II, as well as the general structure of the proposed detector. In Section III, we present the performance measure of interest, the near-far resistance and we show that the near-far resistance of the optimum multiuser detector can be achieved by a linear detector (the decorrelating detector), which is explicitly obtained in Section IV, as well as its implementable version as a linear time-invariant system. Section V gives a numerical comparison of the error probabilities of the decorrelating receiver and the conventional receiver in a scenario of practical interest.

II. Multiuser Communication Model

Let the receiver input signal be

$$r(t) = S(t, b) + n(t) \tag{2.1}$$

where $n(t)$ is white Gaussian noise with power spectral density σ^2 and

$$S(t, b) = \sum_{i=-M}^{M} \sum_{k=1}^{K} b_k(i) \sqrt{w_k(i)} \tilde{s}_k(t - iT - \tau_k) \tag{2.2}$$

is the element of $\mathcal{L}_2$ (the Hilbert space of square-integrable functions) which contains the information sequence $b = \{b(i) = [b_1(i), \cdots, b_K(i)], b_k(i) \in \{-1, 1\}, k = 1, \cdots, K; i =$

Paper approved by the Editor for Spread Spectrum of the IEEE Communications Society. Manuscript received March 24, 1988; revised September 14, 1988. This work was supported in part by the U.S. Army Research Office under Contract DAAL03-87-K-0062. This paper was presented in part at the 1988 IEEE International Symposium on Information Theory, Kobe, Japan, June 1988.

The authors are with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544.

IEEE Log Number 9034846.

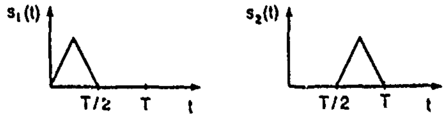


Fig. 1. Example of signature waveforms which can violate the LIA.

$-M,\cdots,M\}$, $\tilde{s}_k(t)$ is the normalized signature waveform of user k and is zero outside the interval $[0,T]$, and $w_k(i)$ is the received energy of user k in the ith time slot. Let $N = 2M+1$ be the length of the transmitted sequence. Without loss of generality it is assumed that the users are numbered such that their delays satisfy $0 \le \tau_1 \le \cdots \le \tau_K < T$. The normalized signal $\tilde{S}(t,b)$ is the receiver input signal corresponding to unit energies.

Define the vector space $L = \{x = [x(-M),\cdots,x(M)] = [[x_1(-M),\cdots,x_K(-M)],\cdots,[x_1(M),\cdots,x_K(M)]]^T,\ x_k(i) \in \mathbb{R},\ k=1,\cdots,K,\ i=-M,\cdots,M\}$, (each element of which can be equivalently viewed as a sequence of N $(K*1)$-vectors or as one single $(NK*1)$-vector), and define the (k,i)th unit vector $u^{k,i}$ in L with components $u_j^{k,i}(l) = \delta_{kj}\delta_{li}$. Note that the set of possible transmitted sequences b is a subset of L, obtained by restricting the components of the vector x to take on the values ± 1. Let $\langle\cdot,\cdot\rangle$ denote the usual inner product on $\mathcal{L}_2$, i.e., the integral of the product over the region of support, with induced norm $\|\cdot\|$. Henceforth, we make the following assumption on $\tilde{S}(t,b)$.

1) Linear Independence Assumption (LIA):

$$\forall \nu \in L,\ \nu \ne 0 \Rightarrow \left\|\tilde{S}(t,\nu)\right\| \ne 0. \tag{2.3}$$

In other words, no matter what the received energies are, the received signal does not vanish everywhere if at least one of the users has transmitted a symbol. This condition fails to hold only in pathological nonpractical cases with very heavy crosscorrelation between the signals, such as the two-user example in Fig. 1. There if the delay between the users is $T/2$, the received signal can be identically zero although transmissions have been made [this happens if, for all i, $b_2(i) = -b_1(i)$]. It is shown in Appendix II that such a situation will arise with probability zero if the *a priori* unknown delays are uniformly distributed, which is the case in the asynchronous channel used by noncooperating users. Basically, in order to violate the LIA, a subset of the users must be effectively synchronous and the modulating signals of this subset have to be heavily correlated. The LIA will be in effect in the rest of the paper. If it is removed all the given results can be generalized in a manner analogous to the treatment of the synchronous transmission case [7].

The sampled output of the normalized matched filter for the ith bit of the kth user, $i=-M,\cdots,M$, is

$$y_k(i) = \int_{iT+\tau_k}^{iT+T+\tau_k} r(t)\tilde{s}_k(t-iT-\tau_k)\,dt \tag{2.4}$$

$$= \int_{-\infty}^{\infty} S(t,b)\tilde{s}_k(t-iT-\tau_k)\,dt + \int_{-\infty}^{\infty} n(t)\tilde{s}_k(t-iT-\tau_k)\,dt \tag{2.5}$$

where the second equality is valid since the signals are zero outside $[0,T]$. It is well established (e.g., [1]) that the whole sequence y of outputs of the bank of K matched filters, with components $y_k(i)$ given by (2.5), for $k=1,\cdots,K$, $i=-M,\cdots,M$, is a sufficient statistic for decision on the most likely transmitted information sequence b. The multiuser demodulation problem which needs to be solved at the receiver is to recover the transmitted sequence $b \in L$ from the sequence $y \in L$. Motivated by the state of the art—where the choice lies between the optimum multiuser detector, which is of exponential complexity and the ad hoc single user detector whose performance degrades to zero for sufficiently high interference energy—we define a class of simple detectors and optimize performance within this class, to obtain an acceptable error probability versus complexity tradeoff.

A *linear detector* for bit i of user k is characterized by $\nu^{k,i} \in L$. The decision of the detector is given by the polarity of the inner product of $\nu^{k,i}$ and the vector y of matched filter outputs, which is equal to

$$\sum_{l=-M}^{M}\sum_{j=1}^{K} \nu_j^{k,i}(l)y_j(l) = \int_{-\infty}^{\infty} \tilde{S}(t,wb)\tilde{S}(t,\nu^{k,i})\,dt + n_{k,i} \tag{2.6}$$

$$= \langle \tilde{S}(t,wb), \tilde{S}(t,\nu^{k,i})\rangle + n_{k,i} \tag{2.7}$$

where for any information sequence b, wb will denote the sequence of amplitudes $wb = \{[\sqrt{w_1(i)}b_1(i),\cdots,\sqrt{w_K(i)}b_K(i)],\ i = -M,\cdots,M\}$. $n_{k,i}$ is the noise component at the output of the cascade of matched filter, sampler and detector, hence is a Gaussian zero-mean random variable with variance given by

$$E[n_{k,j}^2] = \sum_{l,l}\nu_k(l)\nu_j(i)\int_{-\infty}^{\infty}\sigma^2\tilde{s}_k(t-lT-\tau_k)\tilde{s}_j(t-iT-\tau_j)\,dt$$

$$= \sigma^2\left\|\tilde{S}(t,\nu^{k,i})\right\|^2. \tag{2.8}$$

The receiver decides on the ith bit of the kth user according to the rule

$$\hat{b}_k(i) = \operatorname{sgn}\sum_{l=-M}^{M}\sum_{j=1}^{K}\nu_j^{k,i}(l)y_j(l) \tag{2.9}$$

$$= \operatorname{sgn}(\langle \tilde{S}(t,wb), \tilde{S}(t,\nu^{k,i})\rangle + n_{k,i}). \tag{2.10}$$

Wherever it is clear from the context, the superscripts k,i will be omitted.

2) Matrix Notation. It is convenient to introduce the following compact notation. Define the $K*K$ normalized signal crosscorrelation matrices $R(l)$ whose entries are given by

$$R_{kj}(l) = \int_{-\infty}^{\infty}\tilde{s}_k(t-\tau_k)\tilde{s}_j(t+lT-\tau_j)\,dt. \tag{2.11}$$

Then, since the modulating signals are zero outside $[0,T]$

$$R(l) = 0\ \forall\ |l| > 1, \tag{2.12}$$

$$R(-l) = R^T(l), \tag{2.13}$$

and, if the users are numbered according to increasing delays, $R(1)$ is an upper triangular matrix with zero diagonal. Also let $W(l) = \operatorname{diag}([\sqrt{w_1(l)},\cdots,\sqrt{w_K(l)}])$. With this notation the matched filter outputs for $l = \{-M,\cdots,M\}$ can be written in vector form as (cf., [8])

$$y(l) = R(-1)W(l+1)b(l+1) + R(0)W(l)b(l) + R(1)W(l-1)b(l-1) + n(l), \tag{2.14}$$

as can be seen for each component by inserting (2.1) into (2.4). We adopt the convention that $b(-M-1) = b(M+1) = 0$. $n(l)$ is the matched filter output noise vector, with autocorrelation matrix given by

$$E[n(i)n^T(j)] = \sigma^2 R(i-j). \tag{2.15}$$

The entries of the matrices $R(i)$, $i=-1,0,1$ are obtained at the receiver by cross-correlating appropriately delayed replicas of the normalized signature waveforms according to (2.11). Note that no additional complexity is hereby required of the receiver, since knowledge of the normalized signature waveforms and the capability to lock onto the respective delays are necessary for matched filtering and sampling at the instant of maximal signal-to-noise ratio.

In contrast to (2.5) the asynchronous nature of the problem is clearly transparent in (2.14). To make this notation more compact

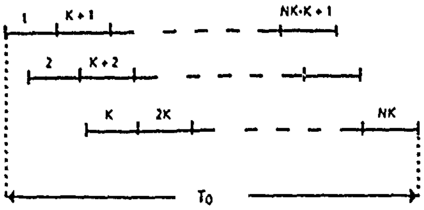


Fig. 2. Equivalent synchronous transmitted sequence.

we define the $NK*NK$ symmetric block-Toeplitz matrix $\mathcal{R}$ and the $NK*NK$ diagonal matrix $\mathcal{W}$, as follows:

$$\mathcal{R} = \begin{pmatrix} R(0) & R(-1) & 0 & \cdots & 0 \\ R(1) & R(0) & R(-1) & & \vdots \\ 0 & R(1) & R(0) & \ddots & 0 \\ \vdots & & \ddots & \ddots & R(-1) \\ 0 & \cdots & 0 & R(1) & R(0) \end{pmatrix}, \quad (2.16)$$

$$\mathcal{W} = \mathrm{diag}\,([\sqrt{w_1(-M)}, \cdots, \sqrt{w_K(-M)}, \cdots, \sqrt{w_1(M)}, \cdots, \sqrt{w_K(M)}]). \quad (2.17)$$

In this notation the matched filter output vector y depends on b via, from (2.14)

$$y = \mathcal{R}\mathcal{W}b + n. \quad (2.19)$$

The matrix $\mathcal{R}$ can be interpreted as the cross-correlation matrix for an equivalent synchronous problem where the whole transmitted sequence is considered to result from $N*K$ users, labeled as shown in Fig. 2, during one transmission interval of duration $T_e = N*T + \tau_K - \tau_1$. Then the results presented here for finite transmission length can be derived via analysis of synchronous multiuser communication, as done in [7]. However, the approach taken in this paper is more general and gives more insight into the nature of the problem. The limit $N \to \infty$ is considered in Section IV-B.

The decision made on the ith bit of the kth user at the output of the detector v is:

$$\hat{b}_k(i) = \mathrm{sgn}\, v^T y = \mathrm{sgn}\, v^T(\mathcal{R}\mathcal{W}b + n). \quad (2.20)$$

As for the inner product, for all x, y in L

$$\langle \tilde{S}(t,x), \tilde{S}(t,y) \rangle = x^T \mathcal{R} y. \quad (2.21)$$

It can be seen from (2 21) and from (2.3) that $\mathcal{R}$ is positive definite.

III. Near-Far Resistance

The main performance measure we are interested in is the bit-error-rate in the high signal-to-background noise region. Thus, even though the background thermal noise is not neglected, our main focus will be on the underlying performance degradation due to multiple-access interference. This performance degradation is conveniently quantified by the *asymptotic efficiency* which was introduced in [1]–[2], and is defined as follows. Let $P_k(\sigma)$ denote the bit-error-rate of the kth user when the spectral level of the background white Gaussian noise is σ^2, and let $e_k(\sigma)$ be such that $P_k(\sigma) = Q(\sqrt{e_k(\sigma)}/\sigma)$.[1]

Then, $e_k(\sigma)$ is actually the energy that the kth user would require to achieve bit-error-rate $P_k(\sigma)$ in the same white Gaussian channel but without interfering users. Hence, we refer to $e_k(\sigma)$ as the *effective energy* of the kth user, and the *efficiency* or ratio between the effective and actual energies $e_k(\sigma)/w_k$ is a number between 0 and 1 which characterizes the performance loss due to the existence of other users in the channel. Thus, the *asymptotic efficiency* (for high SNR) of a transmitter whose bit-error-rate curve and energy are given by $P_k(\sigma)$ and w_k, respectively, is

$$\eta_k = \lim_{\sigma \to 0} \frac{e_k(\sigma)}{w_k} = \sup\left\{0 \le r \le 1;\ \lim_{\sigma \to 0} P_k(\sigma)/Q\left(\frac{\sqrt{r w_k}}{\sigma}\right) < \infty\right\} \quad (3.1)$$

where the last equation follows immediately upon substitution of $P_k(\sigma)$ by its expression in terms of the effective energy. In order to visualize intuitively the asymptotic efficiency, note that the logarithm of the bit-error-rate $P_k(\sigma)$ decays asymptotically with the same slope as the logarithm of the bit-error-rate of a single-user with energy $\eta_k w_k$. Therefore, if $\lim_{\sigma \to 0} P_k(\sigma) > 0$, (i.e., there is an irreducible probability of error even in the absence of background noise), then the asymptotic efficiency is zero. Conversely, nonzero asymptotic efficiency implies that the bit-error-rate goes to zero (as $\sigma \to 0$) exponentially in $1/\sigma^2$.

While asymptotic efficiency and low-noise bit-error-rate are equivalent performance measures, asymptotic efficiency has the advantage of being analytically tractable and of resulting in explicit expressions for the detectors we are interested in. For example, while the probability of error of the optimum multiuser detector does not admit an explicit expression, its asymptotic efficiency is given by [2]

$$\eta_{k,i} = \frac{1}{w_k(i)} \min_{\epsilon \in Z_k} \|\tilde{S}(t, w\epsilon)\|^2 \quad (3.2)$$

where Z_k is the set of error-sequences $\epsilon = \{\epsilon(i) \in \{-1, 0, 1\}^K,\ i = -M, \cdots, M,\ \epsilon_k(i) = 1\}$ that affect the ith bit of the kth user. It was shown in [3] (see also [15]) that the numerical computation of the asymptotic efficiency of optimum multiuser detection given by (3.2) is an NP-complete combinatorial optimization problem.

In an environment where the transmission energies change in time, e.g., if the transmitters are mobile, a performance measure of interest for any detector is its kth *user near-far resistance*, $\overline{\eta_{k,i}}$, which is defined for each detector as its worst case asymptotic efficiency for bit i of user k over all possible energies of the other (interfering and noninterfering) bits, i.e.,

$$\overline{\eta_{k,i}} = \inf_{\substack{w_j(l) \ge 0 \\ (j,l) \ne (k,i)}} \eta_{k,i}. \quad (3.3)$$

In our definition of near-far resistance we model the most general case where the energies of the users are allowed to be time-dependent. This captures the worst case operating conditions of the detector, which are, for example, encountered in mobile radio communication, due to positioning and tracking variations. In the case where the energies are constrained to be arbitrary but nonvarying the present near-far resistance is a lower bound. That case is not amenable to closed-form analysis, since one has to deal with a combinatorial optimization problem.

For illustration consider the two-user case. If the user energies are constant over time, i.e., $w_1(i) = w_1$, $w_2(i) = w_2$, the asymptotic efficiency of the optimal multiuser detector given by (3.2) reduces to [2]:

$$\eta_1 = \min\left\{1,\ 1 + \frac{w_2}{w_1} - 2\max\{|\rho_{12}|, |\rho_{21}|\}\frac{\sqrt{w_2}}{\sqrt{w_1}},\ 1 + 2\frac{w_2}{w_1} - 2(|\rho_{12}| + |\rho_{21}|)\frac{\sqrt{w_2}}{\sqrt{w_1}}\right\}$$

and hence

$$\eta_{\min} \triangleq \min_{\substack{w_2 \\ w_1\ \mathrm{const.}}} \eta_1 = \min\left\{1 - \rho_{12}^2,\ 1 - \rho_{21}^2,\ 1 - \rho_{12}^2 - \rho_{21}^2 + \frac{(|\rho_{12}| - |\rho_{21}|)^2}{2}\right\}, \quad (3.4)$$

[1] $Q(x) = \int_x^{\infty} (1/\sqrt{2\pi}) e^{-v^2/2}\, dv$.

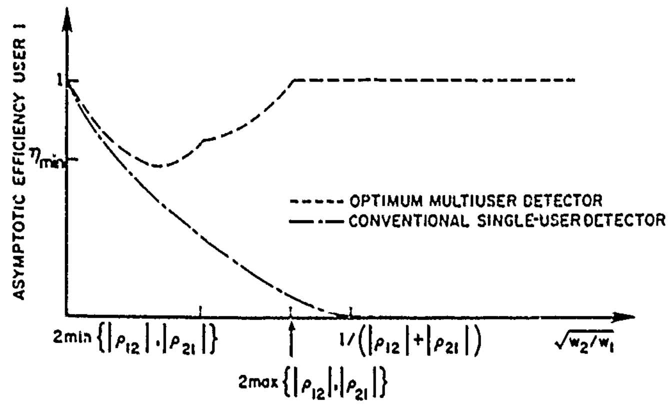


Fig. 3 Asymptotic efficiencies in the two-user case for infinite transmitted sequence length, when the user energies are constant over time (here we chose $|\rho_{12}|, |\rho_{21}| = 0.3, 0.5$).

and analogously for user 2 where $\rho_{12} = R_{12}(0)$ and $\rho_{21} = R_{12}(1)$. The dependence of η_1 for constant energies on the energy ratio is shown in Fig. 3. Note that the optimal multiuser detector is near-far resistant, and in fact has an asymptotic efficiency of unity for sufficiently powerful interference ([2]). Note also that in this case three different error-sequences minimize (3.2) for different values of w_2/w_1, as can be seen from the discontinuity points of the derivative of η. The minimum of η over constant energies, $\eta_{\min}$, is an upper bound on the near-far resistance of optimum multiuser detection $\bar\eta$, which is the minimum asymptotic efficiency over unconstrained energies.

The near-far resistance of the optimal multiuser detector is important since it is the least upper bound on the near-far resistance of any detector, and a measure of the relative performance of any suboptimal detector. From (3.2) and the definition of near-far resistance it is equal to

$$\overline{\eta_{k,i}} = \inf_{\substack{w_j(l)\geq 0\\ (j,l)\neq(k,i)}} \frac{1}{w_k(i)} \min_{\epsilon\in Z_k} \|\tilde{S}(t, w\epsilon)\|^2 \tag{3.5}$$

$$= \inf_{\substack{w_j(l)\geq 0\\ (j,l)\neq(k,i)}} \min_{\epsilon\in Z_k} \left\|\tilde{S}\left(t, \frac{1}{\sqrt{w_k(i)}} w\epsilon\right)\right\|^2 \tag{3.6}$$

$$= \inf_{\substack{y\in L\\ y_k(i)=1}} \|\tilde{S}(t,y)\|^2. \tag{3.7}$$

In Section IV, we obtain a closed-form expression for (3.7) as the reciprocal of the (k, i)th diagonal element (see footnote 2) of the inverse of $\mathcal{R}$. Hence, the near-far resistance of optimum multiuser resistance is guaranteed to be nonzero because of the linear independence assumption of (2.3), which ensures that $\mathcal{R}$ is invertible.

We now turn to the performance analysis of the linear detectors introduced above. The probability of error at decision upon $b_k(i)$ of the linear detector v is, from (2.10):

$$P_k(i) = P(\hat{b}_k(i) \neq b_k(i)) \tag{3.8}$$

$$= P(\langle \tilde{S}(t, wb), \tilde{S}(t, v)\rangle + n_{k,i} < 0 | b_k(i) = 1). \tag{3.9}$$

The equality follows since the hypotheses $+1, -1$ are assumed equally likely. Let B be the set of possible transmitted sequences. From (2.8) $n_{k,i}$ is a zero-mean Gaussian random variable with variance $\sigma^2\|\tilde{S}(t, v)\|^2$, hence the probability of error in (3.9) is a sum of Q-functions, one for each possible interfering bit-combination. For $\sigma \to 0$ the Q-function with the smallest argument dominates the error probability, hence from (3.1), since the expression below can be shown (cf. [15]) to be upper bounded by 1, the asymptotic efficiency achieved by the linear detector v for the ith bit of the kth user is

$$\eta_{k,i}(v) = \frac{1}{w_k(i)} \max{}^2 \left\{0, \min_{\substack{b\in B\\ b_k(i)=1}} \frac{\langle \tilde{S}(t, wb), \tilde{S}(t, v)\rangle}{\|\tilde{S}(t, v)\|}\right\}. \tag{3.10}$$

Knowledge of the asymptotic efficiency of a linear detector is equivalent to knowledge of the worst case probability of error over the bit sequences of the interfering users, since this error probability, which is a Q-function, is set equal to $Q(\sqrt{\eta_{k,i}(v)w_k(i)}/\sigma)$ to obtain (3.10).

For illustration consider the conventional single-user detector in the two-user case. We have $v = u^{k,i}$ (recall that $u^{k,i}$ is the (k, i)th unit vector in the space L of linear detectors). If the user energies are constant over time, i.e., $w_1(i) = w_1$, $w_2(i) = w_2$, the asymptotic efficiency of the conventional single-user detector is found from (3.10) to be

$$\eta_1^c = \max{}^2 \left\{0, 1 - (|\rho_{12}| + |\rho_{21}|)\frac{\sqrt{w_2}}{\sqrt{w_1}}\right\} \tag{3.11}$$

and analogously for user 2. The dependence of η_1^c for constant energies on the energy ratio is shown in Fig. 3. Note that the asymptotic efficiency of the conventional single-user detector is zero for sufficiently high interference energy ($\sqrt{w_2}/\sqrt{w_1} > 1/(|\rho_{12}| + |\rho_{21}|)$). This implies that its near-far resistance is zero, which is what we want to remedy.

There are three quantities of interest in this communication environment, on the one hand the transmitted bit-sequence and the set of energies, both of which depend only on the transmitters and determine the *operating points* for the receiver, and on the other hand the data-processing detector v at the receiver, which we called a *linear detector*. In determining which linear detector to choose at the receiver a useful procedure is the *minimax* approach, in which the design goal is to optimize the worst case performance of the receiver over the class of operating points. Thus we are interested in finding the *maximin linear detector*, whose worst case performance over all allowable input sequences is the highest in the class of linear detectors. The following result quantifies the performance of the maximin detector, in the sequel denoted by v^*.

Proposition 1: There exists a linear detector (which is independent of the received energies) that achieves optimum near-far resistance (i.e., the near-far resistance of the optimum multiuser detector). ●

Proof: From (3.10) the asymptotic efficiency of the linear detector v is

$$\eta_{k,i}(v) = \frac{1}{w_k(i)} \max{}^2 \left\{0, \min_{\substack{b\in B\\ b_k(i)=1}} \frac{\langle \tilde{S}(t, wb), \tilde{S}(t, v)\rangle}{\|\tilde{S}(t, v)\|}\right\} \tag{3.12}$$

$$\times \min_{\substack{b\in B\\ b_k(i)=1}} \frac{1}{w_k(i)} \max{}^2 \left\{0, \frac{\langle \tilde{S}(t, wb), \tilde{S}(t, v)\rangle}{\|\tilde{S}(t, v)\|}\right\} \tag{3.13}$$

$$= \min_{\substack{b\in B\\ b_k(i)=1}} \frac{1}{w_k(i)} \max{}^2 \left\{0, \frac{b^T \mathcal{W}\mathcal{R}v}{\sqrt{v^T\mathcal{R}v}}\right\} \tag{3.14}$$

where in the last equality we have used the compact matrix notation of (2.21) for simplicity. We are interested in the linear detector with the highest worst case asymptotic efficiency, i.e., whose near-far resistance is

$$\overline{\eta_{k,i}(v^*)} = \sup_{\substack{v\in L\\ |\tilde{S}(t,v)|\neq 0}} \inf_{\substack{w_j(l)\geq 0\\ (j,l)\neq(k,i)}} \eta_{k,i}(v) \tag{3.15}$$

$$= \sup_{\substack{v\in L\\ v^T\mathcal{R}v\neq 0}} \underbrace{\inf_{\substack{w_j(l)\geq 0\\ (j,l)\neq(k,i)}} \min_{\substack{b\in B\\ b_k(i)=1}}} \frac{1}{w_k(i)} \max{}^2 \left\{0, \frac{b^T \mathcal{W}\mathcal{R}v}{\sqrt{v^T\mathcal{R}v}}\right\} \tag{3.16}$$

$$= \sup_{\substack{v \in L \\ y^T \mathcal{R} v \neq 0}} \inf_{\substack{y \in L \\ y_k(i)=1}} \max{}^2 \left\{ 0, \frac{y^T \mathcal{R} v}{\sqrt{v^T \mathcal{R} v}} \right\} \tag{3.17}$$

$$= \max{}^2 \left\{ 0, \sup_{\substack{v \in L \\ v^T \mathcal{R} v \neq 0}} \inf_{\substack{y \in L \\ y_k(i)=1}} \frac{y^T \mathcal{R} y}{\sqrt{v^T \mathcal{R} v}} \right\} \tag{3.18}$$

where we have set $y_j(l) = b_j(l)\sqrt{w_j(l)}/\sqrt{w_k(i)}$ for the third equality. Let $M(v, y)$ denote the penalty function $y^T \mathcal{R} v / \sqrt{v^T \mathcal{R} v}$ where the first argument is from the set of detectors and the second from the set of operating points, both specified in (3.18). We show in Appendix 1 that $M(v, y)$ has a saddle point, i.e.,

$$\sup_{\substack{v \in L \\ v^T \mathcal{R} v \neq 0}} \inf_{\substack{v \in L \\ y_k(i)=1}} \frac{y^T \mathcal{R} v}{\sqrt{v^T \mathcal{R} v}} = \inf_{\substack{y \in L \\ y_k(i)=1}} \sup_{\substack{v \in L \\ v^T \mathcal{R} v \neq 0}} \frac{y^T \mathcal{R} v}{\sqrt{v^T \mathcal{R} v}}, \tag{3.19}$$

which establishes the existence of v^* and hence

$$\overline{\eta_{k,i}(v^*)} = \max{}^2 \left\{ 0, \inf_{\substack{y \in L \\ y_k(i)=1}} \sup_{\substack{v \in L \\ v^T \mathcal{R} v \neq 0}} \frac{y^T \mathcal{R} v}{\sqrt{v^T \mathcal{R} v}} \right\} \tag{3.20}$$

$$= \max{}^2 \left\{ 0, \inf_{\substack{y \in L \\ y_k(i)=1}} \sqrt{y^T \mathcal{R} y} \right\} \tag{3.21}$$

$$= \inf_{\substack{y \in L \\ y_k(i)=1}} \|\tilde{S}(t, y)\|^2 \tag{3.22}$$

$$= \overline{\eta_{k,i}} \tag{3.23}$$

where the second equality is obtained in (A.1), the third line follows since $\mathcal{R}$ is nonnegative definite and the last equality was obtained in (3.7). ◇

The reason why the near-far optimum linear receiver achieves the same near-far resistance as the optimum receiver can be understood as follows. Let Ω be the set of multiuser signals modulated by all positive amplitudes, i.e., $\Omega = \{\tilde{S}(t, y),\ y \in L\}$ and let Ξ denote the subset of Ω such that the amplitude of the ith symbol of the kth user is fixed to 1, i.e., $\Xi = \{\tilde{S}(t, y),\ y \in L,\ y_k(i) = 1\}$ (note that Ξ is a convex set, and because of the LIA it does not include the origin). Since the penalty function in (3.18) is invariant to scaling of v and the operator $\mathcal{R}$ is positive definite, (3.18) can be rewritten as

$$\overline{\eta_{k,i}(v^*)} = \max{}^2 \left\{ 0, \sup_{\substack{v \in L \\ |\tilde{S}(t,v)|=1}} \inf_{\substack{y \in L \\ y_k(i)=1}} \langle \tilde{S}(t, y), \tilde{S}(t, v) \rangle \right\} \tag{3.24}$$

$$= \max{}^2 \left\{ 0, \sup_{\substack{v \in \Omega \\ |v|=1}} \inf_{y \in \Xi} \langle y, v \rangle \right\}. \tag{3.25}$$

Therefore the kth user decorrelating filter can be viewed as the unit-norm multiuser waveform whose minimum inner product with the elements of Ξ is highest. But since Ξ is a convex set, that signal is a scaled version of the closest vector in Ξ to the origin (Fig. 4), and its near-far resistance [cf. (3.22)] is the norm squared of that vector. But, as (3.7) indicates, the square of the distance from Ξ to the origin is precisely the near-far resistance of the optimum detector.

Equation (3.7) leads to a nice intuitive interpretation of near-far resistance. Rewrite this equation, using the definition of $\tilde{S}(t, \cdot)$, as

$$\overline{\eta_{k,i}} = \inf_{\substack{y_j(l) \in R \\ (j,l) \neq (k,i)}} \left\| \tilde{s}_k(t - iT - \tau_k) + \sum_{(j,l) \neq (k,i)} y_j(l) \tilde{s}_j(t - lT - \tau_j) \right\|^2 \tag{3.26}$$

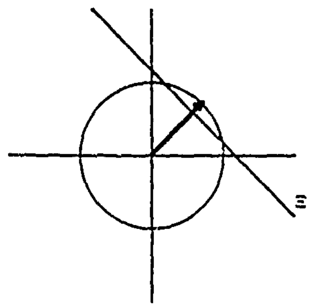


Fig. 4. Interpretation of near-far resistance. Vector in boldface corresponds to decorrelating filter.

Letting $\{y_j(l)\}$ vary over the admissible set, the second term above generates all points of a linear subspace which includes the origin, therefore the infimum in (3.26) is the distance of $\tilde{s}_k(t - iT - \tau_k)$ to this space, i.e.,

$$\overline{\eta_{k,i}} = d^2(\tilde{s}_k(t - iT - \tau_k),\ \text{span}\{\tilde{s}_j(t - lT - \tau_j),\ (j, l) \neq (k, i)\}) \tag{3.27}$$

where $d(a, b)$ denotes the Euclidean distance between the $\mathcal{L}_2$ elements a and b. In the synchronous case because the time-support is disjoint, the infimum in (3.26) is achieved when $y_j(l) = 0,\ l \neq i$, and (3.27) reduces to

$$\overline{\eta_k} = d^2(\tilde{s}_k(t),\ \text{span}\{\tilde{s}_j(t),\ j \neq k\}), \tag{3.28}$$

i.e., the kth user near-far resistance in a synchronous channel is the square of the *distance of the kth user signal to the space spanned by the signals of the interfering users.* Viewing the asynchronous problem in terms of the equivalent synchronous system with $N*K$ users and period NT, the near-far resistance of asynchronous communication allows for the same interpretation. Note, however, that the shifted versions $s_k(t - lT - \tau_k),\ l \neq i$ of the kth user signal affect the near-far resistance of the ith symbol of user k.

The following section characterizes the linear detector that achieves the optimum near-far resistance anticipated by Proposition 1.

IV. The Decorrelating Detector

We first assume N to be finite, as in the case in all communication environments, and characterize the linear filter which achieves the near-far resistance of optimum multiuser detection. This filter is nonstationary for finite N. The limit as $N \to \infty$ is then considered, yielding a stationary noncausal limiting filter, and hence, after appropriate truncation of the noncausal part, an approximation of the near-far optimal linear filter which can be implemented easily

A. The Finite Sequence Length Case

Definition: A decorrelating detector $d^{k,i}$ for the ith bit of the kth user is a linear detector for which

$$\mathcal{R} d^{k,i} = u^{k,i} \tag{4.1}$$

or equivalently, from (2.21), $\langle \tilde{S}(t, v), \tilde{S}(t, d^{k,i}) \rangle = v_k(i)$, for all v in L.

Existence: By the LIA, statement (4.2) below holds for all k, i. Hence, the following equivalences show the existence of the decorrelating detectors for each bit of each user.

$$\forall v \in L \text{ with } v_k(i) \neq 0\text{: } \|\tilde{S}(t, v)\| \neq 0 \tag{4.2}$$

$$\Leftrightarrow \forall v \in L \text{ with } v_k(i) \neq 0\text{: } v^T \mathcal{R} v \neq 0 \tag{4.3}$$

$$\Leftrightarrow \nexists v \in L \text{ with } v_k(i) \neq 0 \text{ s.t. } \mathcal{R} v = 0 \tag{4.4}$$

$\Leftrightarrow$ the $(k, i)^{th}$ column[2] *of* $\mathcal{R}$ is

[2] We refer to the (k, i)th row (or column) of a matrix of the dimension of $\mathcal{R}$ when we want to name the kth row (or column) within the ith block in vertical (horizontal) direction. This notation was adopted since $\mathcal{R}$ is block-Toeplitz.

$$\cdot \quad \times \quad \text{linearly independent of the others} \tag{4.5}$$

$$\Leftrightarrow \exists\, d \text{ s.t. } \mathcal{R}d = u^{k,l}. \tag{4.6}$$

Properties:

i) The decorrelating detector for each bit of each user is invariant with respect to received energies and does not require knowledge thereof.

Proof: Since the elements of the matrix $\mathcal{R}$ are normalized crosscorrelation coefficients, the defining equation (4.1) is energy independent.

ii) The decorrelating detector eliminates the multiuser interference present in the respective matched filter output. (Hence its name).

Proof: From (2.20) the decision made on the ith bit of the kth user at the output of the decorrelating filter d is,

$$\begin{aligned}\hat{b}_k(i) &= \operatorname{sgn}(d^T \mathcal{R} \mathcal{W} b + d^T n)\\ &= \operatorname{sgn}(\sqrt{w_k(i)}\, b_k(i) + d^T n).\end{aligned} \tag{4.7}$$

Interestingly, this natural strategy, though not necessarily optimal for specific user-energies, is optimal with respect to the worst possible distribution of energies.

iii) The kth-user bit-error-rate of the decorrelating detector is independent of the energies of the interfering users $w_j(i)$, $j \neq k$, $i = -M, \cdots, M$.

Proof: It follows from (4.7) that the decision statistic that is compared to a zero threshold is independent of the energies of the interfering users.

iv) The efficiency of the decorrelating detector is independent of the energies and is given by

$$\eta^d_{k,i} = \max{}^2 \left\{ 0, \min_{\substack{b \in B \\ b_k(i)=1}} \frac{1}{\sqrt{w_k(i)}} \frac{\langle \tilde{S}(t, \mathcal{W}b), \tilde{S}(t, d)\rangle}{\|\tilde{S}(t,d)\|} \right\} \tag{4.8}$$

$$= \max{}^2 \left\{ 0, \min_{\substack{b \in B \\ b_k(i)=1}} \frac{1}{\sqrt{w_k(i)}} \frac{\sqrt{w_k(i)}\, b_k(i)}{\sqrt{d_k(i)}} \right\} \tag{4.9}$$

$$= \frac{1}{d_k(i)}, \tag{4.10}$$

which by i) is energy-independent.

v) The decorrelating detector is the worst case optimal linear detector, and achieves the near-far resistance of optimum multiuser detection.

Proof: The proof of Proposition 1 is constructive, hence the first part of v) was obtained as a byproduct in Appendix 1. Here is a shorter proof, using the following fact. Any single linear strategy which is not decorrelating has a near-far resistance of zero. This is shown as follows. The near-far resistance of a linear filter is (cf. (3.18)):

$$\overline{\eta_{k,i}(v)} = \max{}^2 \left\{ 0, \inf_{\substack{v \in L \\ y_k(i)=1}} \frac{v^T \mathcal{R} y}{\sqrt{v^T \mathcal{R} y}} \right\}. \tag{4.11}$$

Unless $\mathcal{R}v = u^{k,i}$ (note invariance of η to scaling of v) the value of the inf-term is $-\infty$. Hence any linear filter which is not decorrelating has a near-far resistance $\bar{\eta} = 0$. This fact together with the nonzero asymptotic efficiency (4.10) of the decorrelating detector establish optimality of the decorrelating detector within the class of linear filters. Therefore the second part of v) results from Proposition 1.

Note that since the asymptotic efficiency of the decorrelating detector is independent of energies (Property iv) it equals the near-far resistance. This gives us an explicit solution for the Hilbert space optimization problem we obtained for the near-far resistance of optimal multiuser detection in (3.7), namely,

$$\overline{\eta_{k,i}} = \eta^d_{k,i} = \frac{1}{d_k(i)} \tag{4.12}$$

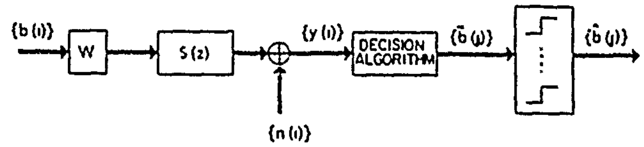


Fig. 5. Equivalent communication system.

and outlines an alternative proof for Proposition 1: we could have explicitly solved the above optimization problem by proceeding along the same lines as in Appendix 1, postulated the decorrelating detector by reasoning as in Fact under v), and shown that the asymptotic efficiency of the decorrelating detector and the near-far resistance of optimal multiuser detection are equal (see [7]). However, the game theoretic proof provides more insight into the nature of the solution.

Property iii) is of special importance. By this property the decorrelating detector does not become multiple-access limited, no matter how strong the multiple-access interference is. Also the decorrelating detector demodulates the data perfectly in the absence of noise, as can be seen from (4.7).

Characterization: We would now like to find an explicit expression for the decorrelating detector which we have up to now defined implicitly. It follows immediately from (4.1) and the uniqueness of the inverse of an invertible matrix that the decorrelating detector for the ith bit of user k is the (k, i)th row of the inverse of $\mathcal{R}$.

From the above and (4.10) the asymptotic efficiency of the decorrelating detector for the ith bit of user k is given by the (k, i)th diagonal element of the inverse of $\mathcal{R}$:

$$\eta^d_{k,i} = \frac{1}{\mathcal{R}^{-1}_{(k,i),(k,i)}}. \tag{4.13}$$

For the values of N encountered in practical applications, inverting a $NK{*}NK$ matrix is not possible. This issue is addressed in Section IV-B where we represent the decorrelating detector as a K-input K-output time-varying linear filter, and then show that in the limit as N tends to infinity the filter becomes time-invariant.

B. The Limiting Case $N \to \infty$

Proposition 2: As the length of the transmitted sequence increases ($N \to \infty$) the decorrelating detector approaches the K-input K-output linear time-invariant filter with transfer function

$$G(z) = [R^T(1)z + R(0) + R(1)z^{-1}]^{-1}. \tag{4.14}$$

Proof: From (2.14) and (2.13) the matched filter outputs for $l = \{-M, \cdots, M\}$ are

$$\begin{aligned}y(l) = {}& R^T(1)W(l+1)b(l+1) + R(0)W(l)b(l)\\ &+ R(1)W(l-1)b(l-1) + n(l)\end{aligned} \tag{4.15}$$

where $b(-M-1) = b(M+1) = 0$. Taking z-transforms and letting N go to infinity we have

$$Y(z) = S(z)[WB](z) + N(z) \tag{4.16}$$

where $[WB](z)$ is the z-transform of the sequence $wb = \{[\sqrt{w_1(i)}\, b_1(i), \cdots, \sqrt{w_K(i)}\, b_K(i)]\}$, the matrix $S(z)$ is

$$S(z) = R^T(1)z + R(0) + R(1)z^{-1} \tag{4.17}$$

and $Y(z)$, $B(z)$ and $N(z)$ are, respectively, the vector-valued z-transforms of the matched filter output sequence, the transmitted sequence, and the noise sequence at the output of the matched filters. $S(z)$ can be interpreted as the equivalent transfer function of the multiuser communication system between transmitter and decision algorithm, as illustrated in Fig. 5. In this setting the optimal receiver problem is to find the transfer function matrix $G(z)$ of a K-input K-output linear time-invariant filter, at the output of which

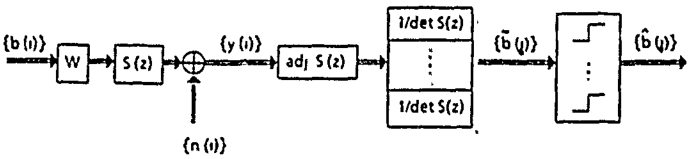


Fig. 6. Interpretation of the decorrelating detector.

a sign-decision yields estimates of the transmitted sequence which are optimal in a certain sense. In our case the optimality criterion is the near-far resistance, and we have demonstrated that the optimal filter is the decorrelating filter, which is the filter that eliminates the multiuser interference, i.e., is the K-input K-output time invariant linear filter which recovers the transmitted data in the absence of noise. Its transfer function is therefore the inverse of the equivalent transfer function $S(z)$:

$$G(z) = [S(z)]^{-1}. \tag{4.18}$$

◇

The effect of the inverse filter $[S(z)]^{-1}$ can be interpreted as illustrated in Fig. 6. The decorrelating filter can be viewed as the cascade of a finite impulse response filter with transfer function *adjoint* $S(z)$, which decorrelates the users, but introduces intersymbol interference among the previously noninterfering symbols of the same user, and of a second filter, consisting of a bank of K identical filters with transfer function $[\det S(z)]^{-1}$, which removes this intersymbol interference. Whereas the region of convergence of the z-transform can always be chosen so as to make $S(z)$ invertible, attention has to be paid to the issue of stability.

Proposition 3: There is a stable, noncausal realization of the decorrelating detector, if and only if the signal cross-correlations are such that

$$\det S(e^{j\omega}) = \det [R^T(1)e^{j\omega} + R(0) + R(1)e^{-j\omega}] \neq 0, \ \forall\, \omega \in [0, 2\pi]. \tag{4.19}$$

●

Proof: As long as $\det S(z)$ has no zeros on the unit circle, a nonempty convergence region of $S^{-1}(z)$ can be chosen which includes the unit circle. Thus, stability can be achieved. But, since $R(0)$ is symmetric,

$$\det S(z) = \det S^T(z) = \det S(z^{-1}).$$

Hence, the stable version of the decorrelating detector will be noncausal. (As a side remark, the matrix $S(e^{j\omega})$ is nonnegative define for all ω, cf. [15]). ◇

Condition (4.19) is equivalent to the limit of the LIA as $N \to \infty$. Both are necessary and sufficient conditions for system invertibility. The LIA requires that the output of a system (the system between the user bit-streams and the matched filter outputs) not be identically zero if the input is nonzero. Hence different inputs generate different outputs, i.e., the system is invertible. For a linear system the requirement that nonzero input produce nonzero output is equivalent to requiring that the transfer matrix be nonsingular on the unit circle. Assume the transfer matrix is singular at the angular frequency ω_0. Necessity follows since otherwise the input sequence consisting of a complex exponential at ω_0 times a vector in the nullspace of the transfer matrix evaluated at ω_0 yields zero output; since the transfer function on the unit circle gives the magnitude and phase of the system response to complex exponentials. On the other hand, sufficiency can be established by using Parseval's relation extended to multivariable systems:

$$\sum_{n=-\infty}^{\infty} \|y_n\|^2 = \frac{1}{2\pi}\int_0^{2\pi} \|Y(e^{j\omega})\|^2\, d\omega = \frac{1}{2\pi}\int_0^{2\pi} \|H(e^{j\omega})X(e^{j\omega})\|^2\, d\omega.$$

Hence, for a zero output sequence y_n the vector $H(e^{j\omega})X(e^{j\omega})$ has to vanish for all ω, which implies that $H(e^{j\omega})$ is singular whenever $X(e^{j\omega})$ is nonzero. This establishes the claimed equivalence.

Proposition 4: Condition (4.19) of Proposition 3 is equivalent to

$$\min_{\substack{x^* x = 1 \\ x \in \mathbb{C}}} \left(x^* R(0) x - \sqrt{(x^* R_+ x)^2 + (x^* R_- x)^2}\right) > 0 \tag{4.20}$$

where $R_+ = R^T(1) + R(1)$ and $R_- = j(R^T(1) - R(1))$. The * denotes the complex conjugate. ●

Note that both R_+ and R_- are Hermitian. The proof of Proposition 4 is given in [15], together with the following two results.

—A necessary condition for (4.20) is that the matrices $R(0) + R(1) + R^T(1)$ and $R(0) - R(1) - R^T(1)$ be nonsingular.

—A sufficient condition for (4.20) is that

$$\lambda^2_{\min}(R(0)) > \max\{\lambda^2_{\max}(R_+), \lambda^2_{\min}(R_+)\} + \lambda^2_{\max}(R_-).$$

The following results quantify the asymptotic efficiency achieved by the limiting decorrelating detector.

Proposition 5: Let

$$[S(z)]^{-1} = \sum_{m=-\infty}^{\infty} D(m) z^{-m}. \tag{4.21}$$

Then the asymptotic efficiency of the limiting decorrelating detector for the kth user is given by

$$\eta_k^d = \frac{1}{D_{kk}(0)} \tag{4.22}$$

$$= \left[\frac{1}{2\pi}\int_0^{2\pi} [R^T(1)e^{j\omega} + R(0) + R(1)e^{-j\omega}]^{-1}_{kk}\, d\omega\right]^{-1}. \tag{4.23}$$

●

Proof: From Proposition 2 the z-transform of the decision statistic at the output of the limiting decorrelating detector is given by

$$G(z)Y(z) = [WB](z) + [S(z)]^{-1}N(z) = [WB](z) + N'(z)$$

where $N'(z)$ is the z-transform of the (stationary) filtered Gaussian background noise vector sequence. The z-transform of its covariance matrix sequence $E[n'(\cdot)n'^T(\cdot + i)]$ is equal to $\sigma^2[S(z)]^{-1}$, hence with (4.21) n'_k is a zero-mean Gaussian random variable with variance $\sigma^2 D_{kk}(0)$. Therefore, the probability of error for the kth user equals

$$P_k = P(n'_k > \sqrt{w_k}) = Q\left(\frac{\sqrt{w_k}}{\sigma\sqrt{D_{kk}(0)}}\right). \tag{4.24}$$

From here, using the definition of asymptotic efficiency, the first equality follows. For the second, note that applying the inverse z-transform and definition (4.21), we obtain

$$D_{kk}(0) = \frac{1}{2\pi}\int_0^{2\pi} [S(e^{j\omega})]^{-1}_{kk}\, d\omega,$$

and the result follows using (4.17). ◇

Proposition 6: The asymptotic efficiency of the limiting decorrelating detector for the kth user is strictly positive, and lower bounded by

$$\eta_k^d \geq \left[\max_{\omega \in [0, 2\pi]} \left|[R^T(1)e^{j\omega} + R(0) + R(1)e^{-j\omega}]^{-1}_{kk}\right|\right]^{-1} > 0. \tag{4.25}$$

●

Proof: From (4.22), (4.23)

$$D_{kk}(0) \le \max_{\omega \in [0, 2\pi]} \left| [R^T(1)e^{j\omega} + R(0) + R(1)e^{-j\omega}]^{-1}_{kk} \right|. \quad (4.26)$$

Hence,

$$\eta_k^d = \frac{1}{D_{kk}(0)} \ge \left[\max_{\omega \in [0, 2\pi]} \left| [R^T(1)e^{j\omega} + R(0) + R(1)e^{-j\omega}]^{-1}_{kk} \right| \right]^{-1}$$
$$\ge \frac{\min_{\omega} |\det [R^T(1)e^{j\omega} + R(0) + R(1)e^{-j\omega}]|}{\max_{\omega} |\mathrm{adj}_k\, [R^T(1)e^{j\omega} + R(0) + R(1)e^{-j\omega}]|}, \quad (4.27)$$

which is positive by Proposition 3. ◇

Proposition 7: In the two-user case let $R_{12}(0) = \rho_{12}$ and $R_{12}(1) = \rho_{21}$. Then the asymptotic efficiency of the decorrelating detector for infinite sequence length is given by

$$\eta_1^d = \eta_2^d = \sqrt{(1 - \rho_{12}^2 - \rho_{21}^2)^2 - 4\rho_{12}^2\rho_{21}^2}$$
$$= \sqrt{[1 - (\rho_{12} + \rho_{21})^2][1 - (\rho_{12} - \rho_{21})^2]}. \quad (4.28)$$

●

Proof: This formula can be obtained by particularizing Proposition 5 or by minimizing the asymptotic efficiency of optimal multiuser detection in the two-user case with respect to energies. Alternatively, we will prove (4.28) by taking the limit as $N \to \infty$ of the asymptotic efficiency of the decorrelating filter for the central bits in a length N sequence. We will then have proved that in the two-user case the limit of the asymptotic efficiency of the finite-length decorrelating detector as $N \to \infty$ is indeed the asymptotic efficiency of the limiting decorrelating detector.

Recall that the asymptotic efficiency of the decorrelating detector is given by the reciprocal of the corresponding diagonal element of $\mathcal{R}^{-1}$. We need to find explicit expressions for the central diagonal elements of the inverse of the matrix $\mathcal{R}$ as a function of N. We have

$$\mathcal{R} = \begin{pmatrix} 1 & \rho_{12} & 0 & 0 & \\ \rho_{12} & 1 & \rho_{21} & 0 & \\ 0 & \rho_{21} & 1 & \rho_{12} & \ddots \\ 0 & 0 & \rho_{12} & 1 & \ddots \\ & \ddots & \ddots & \ddots & \ddots \end{pmatrix}. \quad (4.29)$$

Denote by Δ_n the determinant of the above $n*n$ matrix. It is easy to see from the structure of $\mathcal{R}$ that Δ_n satisfies the recursion

$$\Delta_n = \Delta_{n-1} - \begin{cases} \rho_{12}^2\, \Delta_{n-2}, & n \text{ even} \\ \rho_{21}^2\, \Delta_{n-2}, & n \text{ odd}. \end{cases} \quad (4.30)$$

Hence, we can write

$$\begin{bmatrix} \Delta_{2n} \\ \Delta_{2n-1} \end{bmatrix} = \begin{bmatrix} 1 - \rho_{12}^2 & -\rho_{21}^2 \\ 1 & -\rho_{21}^2 \end{bmatrix} \begin{bmatrix} \Delta_{2n-2} \\ \Delta_{2n-3} \end{bmatrix}. \quad (4.31)$$

If we consider the sequence of $4n*4n$ matrices for simplicity, the central diagonal element of the inverse of $\mathcal{R}$ is $\Delta_{4n}/(\Delta_{2n-1}\Delta_{2n})$. Hence, after introducing the state vector

$$x_n = \begin{bmatrix} \Delta_{2n} \\ \Delta_{2n-1} \end{bmatrix}, \quad (4.32)$$

we see that finding Δ_{2n}, Δ_{2n-1} requires finding the trajectory of the unforced linear dynamic system

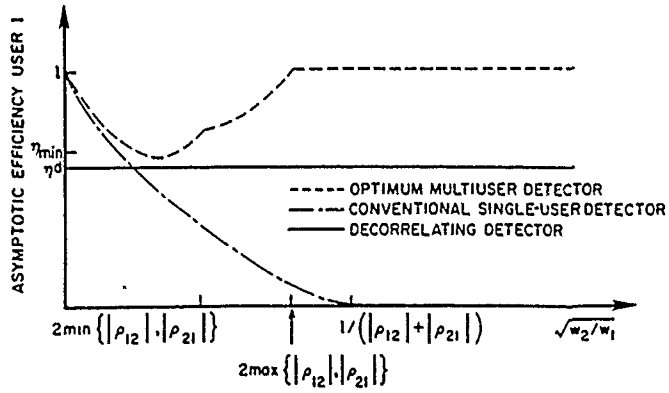


Fig. 7. Asymptotic efficiencies in the two-user case for infinite transmitted sequence length, when the user energies are constant over time (here we chose $|\rho_{12}|$, $|\rho_{21}| = 0.3, 0.5$ which yields $\eta_{\min} = 0.68$, $\eta^d = 0.59$)

$$x_n = \begin{bmatrix} 1 - \rho_{12}^2 & -\rho_{21}^2 \\ 1 & -\rho_{21}^2 \end{bmatrix} x_{n-1}, \; x_1 = \begin{bmatrix} 1 - \rho_{12}^2 \\ 1 \end{bmatrix},$$

i.e.,

$$x = \begin{bmatrix} 1 - \rho_{12}^2 & -\rho_{21}^2 \\ 1 & -\rho_{21}^2 \end{bmatrix}^n \begin{bmatrix} 1 \\ 0 \end{bmatrix}. \quad (4.33)$$

The eigenvalues of this system are found to be

$$\lambda_{1,2} = \frac{1 - \rho_{12}^2 - \rho_{21}^2 \pm \sqrt{(1 - \rho_{12}^2 - \rho_{21}^2)^2 - 4\rho_{12}^2\rho_{21}^2}}{2}.$$

We see $0 < \lambda_1 < \lambda_2 < 1$. After finding the corresponding eigenvectors it follows that:

$$x_n = \begin{bmatrix} \lambda_1 + \rho_{12}^2 & \lambda_2 + \rho_{21}^2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{bmatrix} \cdot \begin{bmatrix} 1 & -(\lambda_2 + \rho_{21}^2) \\ -1 & \lambda_1 + \rho_{21}^2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \frac{1}{\lambda_1 - \lambda_2}$$
$$= \begin{bmatrix} \lambda_1 + \rho_{12}^2 & \lambda_2 + \rho_{21}^2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \lambda_1^n \\ -\lambda_2^n \end{bmatrix} \frac{1}{\lambda_1 - \lambda_2}. \quad (4.34)$$

Hence the central diagonal element of the inverse of $\mathcal{R}$ is

$$\frac{\Delta_{4n}}{\Delta_{2n-1}\Delta_{2n}} = \frac{[1 \;\; 0]x_{2n}}{[0 \;\; 1]x_n[1 \;\; 0]x_n}$$
$$= \frac{(\lambda_1/\lambda_2)^{2n}(\lambda_1 + \rho_{21}^2) - (\lambda_2 + \rho_{21}^2)}{[(\lambda_1/\lambda_2)^n - 1][(\lambda_1/\lambda_2)^n(\lambda_1 + \rho_{21}^2) - (\lambda_2 + \rho_{21}^2)]} \cdot (\lambda_1 - \lambda_2). \quad (4.35)$$

So finally

$$\eta^d = \lim_{n \to \infty} \frac{\Delta_{4n}}{\Delta_{2n-1}\Delta_{2n}} = \lambda_2 - \lambda_1 = \sqrt{(1 - \rho_{12}^2 - \rho_{21}^2)^2 - 4\rho_{12}^2\rho_{21}^2}.$$

◇

Fig. 7 shows the asymptotic efficiency of the decorrelating detector for infinite transmitted sequence length in the two user case. Note its invariance with respect to energies. The discrepancy between η^d and $\eta_{\min}$, defined in (3.4), is due to the fact that $\eta_{\min}$ is higher than the near–far resistance of optimum multiuser detection, since for $\eta_{\min}$ the energies are constrained to be constant over time.

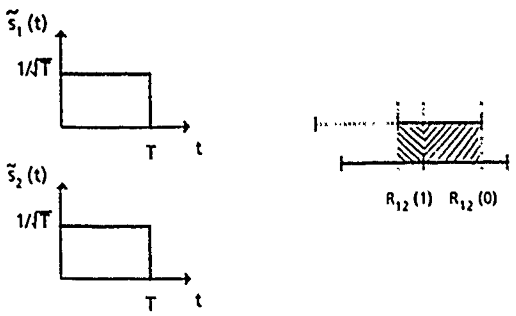


Fig. 8. Signals and crosscorrelations of example (4.42).

The fact that the stable version of the decorrelating filter turns out to be noncausal is not surprising. Due to the lack of synchronism among the users any decision based on less than the entire received waveform is suboptimal. In practice, since the filter is stable, the more remote symbols will count less heavily, and truncation of the noncausal part will be performed after a suitable delay without affecting performance appreciably. For illustration consider the two-user case where we let $R_{12}(0) = \rho_{12}$ and $R_{12}(1) = \rho_{21}$. Then

$$S(z) = \begin{pmatrix} 1 & \rho_{12} + \rho_{21} z^{-1} \\ \rho_{12} + \rho_{21} z & 1 \end{pmatrix}$$

and the transfer function of the decorrelating detector as given by (4.27) is

$$S^{-1}(z) = \frac{1}{1 - \rho_{12}^2 - \rho_{21}^2 - \rho_{12}\rho_{21} z - \rho_{12}\rho_{21} z^{-1}} \cdot \begin{pmatrix} 1 & -(\rho_{12} + \rho_{21} z^{-1}) \\ -(\rho_{12} + \rho_{21} z) & 1 \end{pmatrix}. \quad (4.36)$$

We are interested in the impulse response $f(n)$ of the IIR part of the above filter. Taking the inverse z-transform it is found to be

$$f(n) = Z^{-1}\left[\frac{1}{1 - \rho_{12}^2 - \rho_{21}^2 - \rho_{12}\rho_{21} z - \rho_{12}\rho_{21} z^{-1}}\right] = \frac{\xi^{|n|}}{\eta} \quad (4.37)$$

where $\xi = (1 - \rho_{12}^2 - \rho_{21}^2 - \eta)/(2\rho_{12}\rho_{21})$ and η is the asymptotic efficiency which is given by Proposition 7. It can be checked that $|\xi| \leq 1$, with equality if $|\rho_{12}| + |\rho_{21}| = 1$, which can be shown to coincide with the condition imposed by Proposition 3 for the two-user case. In the latter case the asymptotic efficiency is zero, which follows from Proposition 7. Otherwise, since $|\xi| < 1$ the limiting filter is stable, with symmetric coefficients which decay with rate ξ. In practical applications the filter will be approximated up to any desired precision by truncation of the noncausal part to a finite number of filter coefficients. For illustration the decay rate ξ of the filter coefficients and the achievable asymptotic efficiency η are plotted in Fig. 9 as functions of ρ_{12} and ρ_{21}.

Poor cross-correlation properties among the signature waveforms could imply that the limiting filter $G(z)$ does not exist, although the decorrelating detector exists for finite-length transmitted sequences. We give an example to illustrate this fact. For $K = 2$ it is straightforward to show that the condition of Proposition 3 is satisfied for all signal constellations for which $|R_{12}(0)| = |R_{12}(1)| \neq 1$. This is the case unless the normalized waveforms coincide modulo circular shifts and sign changes.

Consider the trivial signal case where both users are assigned the same rectangular waveform, as shown in Fig. 8. Abbreviate $R_{12}(0)$, which is the crosscorrelation between bits in the same signaling interval, by $r = \tau/T \in [0, 1)$, then in this case $R_{12}(1)$, which is the crosscorrelation between bits in adjacent intervals, is $1 - r$. Then,

$$S(z) = \begin{pmatrix} 1 & r + (1-r)z^{-1} \\ r + (1-r)z & 1 \end{pmatrix} \quad (4.38)$$

becomes singular for $z = 1$, hence there is no stable limiting inverse filter. And if it existed its asymptotic efficiency, as given by (4.28), would be zero. This is not surprising, for an infinite sequence of transmitted bits where both users use the same waveform. However, for finite length sequences advantage can be taken of the marginal effects of having bits which are not affected by either past or future bits. For finite N the decorrelating detector exists unless $r = 0$, i.e., when the transmissions are not synchronous. This is in accord with the multiarrival condition given in Appendix 2, and with the results obtained in the synchronous case [7].

V. Error Probabilities: Numerical Examples

In the sequel, we compare the performances of the conventional and of the decorrelating detector. Without loss of generality we consider the error probability of user 1 in a channel shared by several active users. The conventional detector decides for the sign of the kth component of the matched filter output vector, given by (2.14). Therefore its average error probability over the bit sequences of the interfering users equals

$$\frac{1}{2^2(K-1)} \sum_{b_j(0),\, b_j(-1),\, j \neq 1} Q\left(\frac{\sqrt{w_1} - \sum_{j=2}^{K} [R_{1j}(0) b_j(0) + R_{1j}(1) b_j(-1)] \sqrt{w_j}}{\sigma}\right), \quad (5.1)$$

whereas its worst case error probability over the interfering bit sequences equals

$$Q\left(\frac{\sqrt{w_1} - \sum_{j=2}^{K} [|R_{1j}(0)| + |R_{1j}(1)|] \sqrt{w_j}}{\sigma}\right). \quad (5.2)$$

The probability of error of the decorrelating detector equals, from (4.24),

$$Q\left(\frac{\sqrt{w_1}}{\sigma\sqrt{D_{11}(0)}}\right),$$

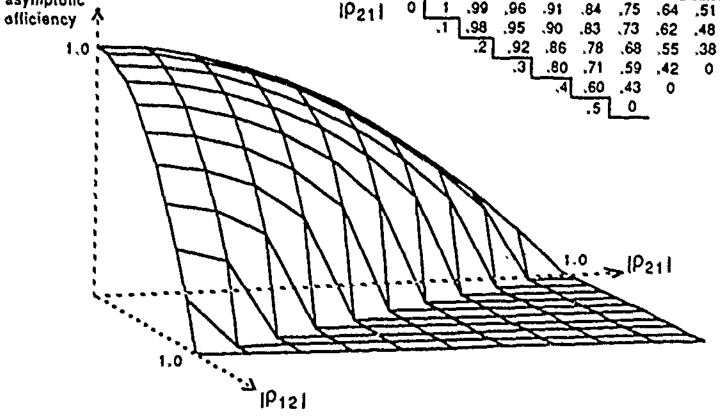
with (the equivalence with (4.23) is easy to show, cf. [15])

$$D_{11}(0) = \frac{1}{\pi}\int_0^{\pi} [R(1)^T e^{j\omega} + R(0) + R(1) e^{-j\omega}]_{11}^{-1}\, d\omega. \quad (5.3)$$

The delays enter the above formulas implicitly via the crosscorrelation matrices, which are functions thereof and of the chosen signature sequences. In the following examples, we have chosen a set of spread-spectrum m-sequences of length 31.

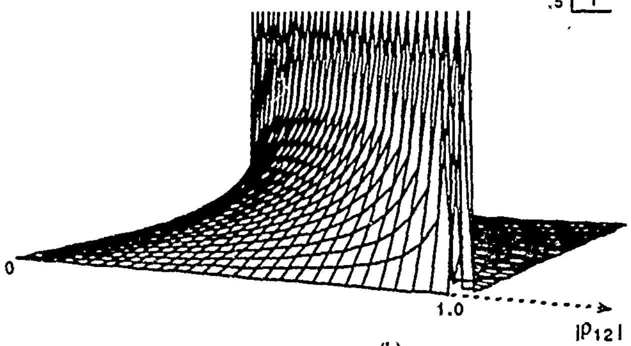
In Fig. 10 we use, for comparison purposes to previous works ([14], [1]), the set of 3 sequences reported in [12, Table V] to be optimal with respect to a signal-to-multiple-access interference parameter when the conventional detector is used. We consider a baseband environment with $K - 1$ active equal energy interferers, whose delay relative to each other is fixed. Fig. 10, for $K = 3$, shows the 1st user error probability of the conventional receiver versus SNR_1, the signal-to-background-noise ratio of user 1, for different values

$\lvert\rho_{21}\rvert$ \ $\lvert\rho_{12}\rvert$	0	.1	.2	.3	4	.5	.6	.7	.8	.9	1
0	1	.99	.96	.91	.84	.75	.64	.51	.36	.19	0
.1		.98	.95	.90	.83	.73	.62	.48	.31	0	
.2			.92	.86	.78	.68	.55	.38	0		
.3				.80	.71	.59	.42	0			
.4					.60	.43	0				
.5						0					



(a)

$\lvert\rho_{21}\rvert$ \ $\lvert\rho_{12}\rvert$	0	.1	.2	.3	.4	.5	.6	.7	.8	.9	1
0	0	0	0	0	0	0	0	0	0	0	0
.1		.01	.02	.03	.05	.07	.1	.14	.24	1	
.2			.04	.07	.1	.14	.21	.33	1		
.3				.11	.16	.24	.37	1			
.4					.25	.39	1				
.5						1					



(b)

Fig. 9. (a) Asymptotic efficiency of the decorrelating detector for two users as a function of the partial crosscorrelations of their signature waveforms. (b) Decay rates of the coefficients of the IIR part of the decorrelating detector for two users, symmetric in ρ_{12} and ρ_{21}.

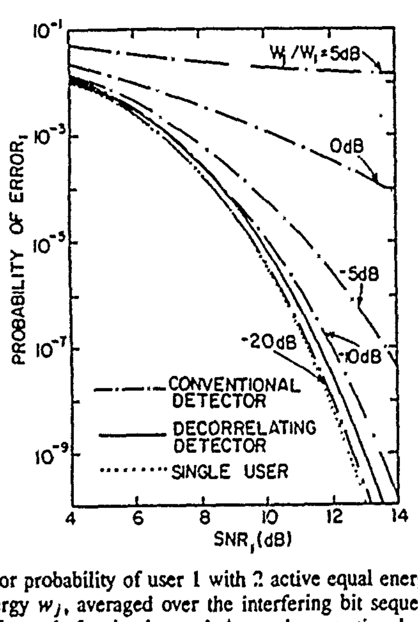


Fig 10. Error probability of user 1 with 2 active equal energy interferers, each of energy w_j, averaged over the interfering bit sequences and over the delay of user 1, for the decorrelating and conventional receiver versus the SNR of user 1, for m-sequences of length 31 and different interference levels.

of the energy ratio SNR_j/SNR_1, averaged over the bit sequences of the two interferers and over the delay of user 1. Also shown are the 1st user error probability of the decorrelating detector and the error probability of the single user channel. From Fig. 10 we see the strong dependence of the performance of the conventional receiver on the relative energies of the active users. While the error probability of the decorrelating detector is invariant to the energy of interfering users, the performance of the conventional receiver deteriorates rapidly for increasing interference, till for an energy ratio above 5 dB the conventional receiver becomes practically multiple-access limited. (For a sufficiently high level of nonorthogonal interference the error probability of the conventional receiver can be seen to become irreducible. E.g., in the two-user synchronous case, for $\sqrt{w_2}/\sqrt{w_1} = (1+\Delta)/\rho$ where ρ is the normalized crosscorrelation coefficient between the two signature signals and $\Delta \geq 0$, the error probability of the conventional receiver tends to 1/4 if $\Delta = 0$ and to 1/2 if $\Delta > 0$ for increasing SNR of user 1). Note that if the energies of all the users are equal the decorrelating detector is around two orders of magnitude better than the conventional receiver at 10 dB. Only if the multiple-access interference level plays a subordinate role compared to the background noise does the conventional detector outperform the decorrelating detector, which pays a penalty for combatting the interference instead of ignoring it. Similar results were obtained regardless of the actual value of the relative delay between the two interfering users.

Fig. 11 shows the same setting as above, in the case $K = 6$. We have used the set of autooptimal m-sequences of length 31 found in

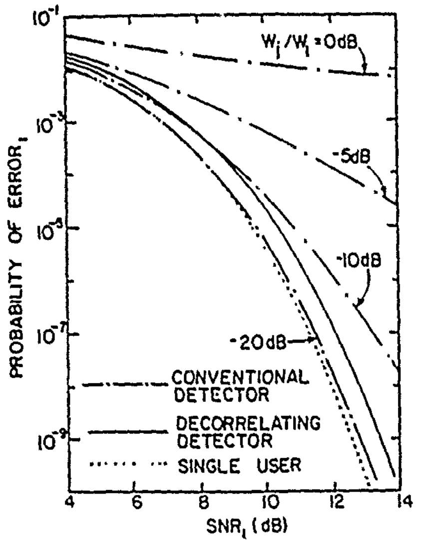


Fig. 11. Same as Fig. 10, with 5 active equal energy interferers.

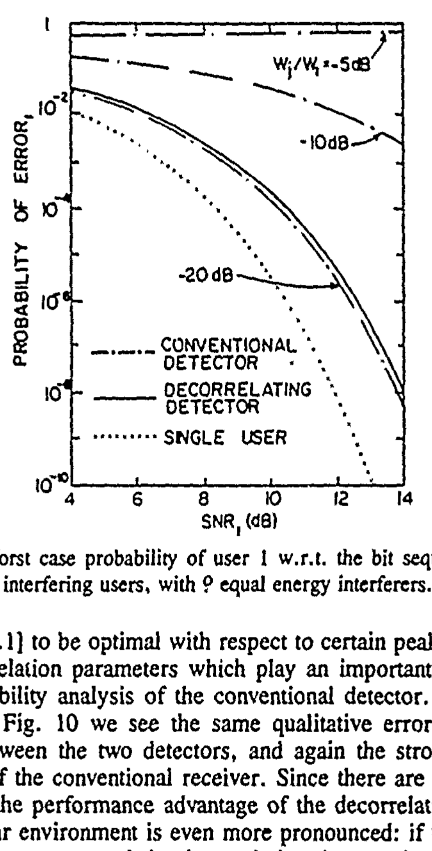


Fig. 12. Worst case probability of user 1 w.r.t. the bit sequences of the interfering users, with 9 equal energy interferers.

[13, Fig. A.1] to be optimal with respect to certain peak and mean-square correlation parameters which play an important role in the error probability analysis of the conventional detector. Comparing Fig. 11 to Fig. 10 we see the same qualitative error probability relation between the two detectors, and again the strong near-far limitation of the conventional receiver. Since there are more active interferers the performance advantage of the decorrelating detector in a near-far environment is even more pronounced: if the energies of all the users are equal the decorrelating detector is almost three orders of magnitude better than the conventional receiver at 10 dB.

Finally, Fig. 12 shows the worst case probability of the conventional detector over the sequences of interfering users, as given by (5.2), for $K = 10$. The signature sequence set used for $K = 6$ has been expanded—without trying to optimize, as before, with respect to the performance of the conventional detector. The shown error probabilities are typical, varying very little if different sets of delays are used because of the good crosscorrelation properties of m-sequences.

Overall the generated error probability curves show the pronounced superiority of the decorrelating receiver in a near-far environment, and whenever sufficiently many users are active even if their energies are well below the energy of the desired user. Note, finally, that the selected signature sequences were optimal with respect to the performance of the conventional receiver. It would be interesting to investigate the possible performance gain of using the decorrelating detector in conjunction with a set of signature sequences optimized for its use.

VI. Conclusions

In this paper, we have obtained a linear multiuser detector, the decorrelating detector, for demodulation of asynchronous code-division multiplexed signals in white Gaussian channels. The bit-error-rate of this detector is independent of the energy of the interfering users and exhibits the same degree of near-far resistance as the optimum multiuser detector obtained in [1]. Since the decorrelating detector does not require knowledge of the received energies and its complexity is only linear in the number of users, it emerges as the solution of choice in near-far environments with a large number of users.

In applications where each receiver is interested in demodulating the information transmitted by only one user, it is easy to decentralize the K-user decorrelating receiver since it can be implemented as K separate (continuous-time) single-input (discrete-time) single-output filters. Each of those filters can be viewed as a modification of the conventional single-user matched filter where instead of correlating the channel output with the signature waveform of the user of interest, we use its projection on the subspace orthogonal to the space spanned by the interfering signals.

Note, finally, that if the filter is actually an approximation to the decorrelating receiver, due to, for example, finite accuracy in the computation of the crosscorrelations or truncation of the impulse response (Section IV-B), it will no longer be orthogonal to the subspace of the interfering signals and therefore it will not be near-far resistant in the worst case sense adopted in this paper. However, the effect on the bit-error-rate will be arbitrarily small with a good enough approximation to the decorrelating receiver, and therefore the bit-error-rate will be very insensitive to the energy level of the interferers. Hence, the resistance to the near-far problem can be preserved within any desired energy range.

Appendix I

Saddle-Point Property in (3.19)

Though the penalty function of (3.18) looks similar to the signal-to-noise ratio functional encountered in the robust matched filtering problem [5], $|\langle h, s\rangle|^2/\langle h, \sum h\rangle$, the problem is different here because the numerator can be negative. Thus we have to establish the result "from scratch." In order to show that $M(\nu, y)$ has a saddle point, i.e., satisfies (3.19), we show that it satisfies the requirements of the following theorem.

Theorem [4, Thm. 2.1]: Suppose Q is a convex set and $M(\nu, \cdot)$ is convex on Q for every $\nu \in H$. Then if (ν_L, y_L) is a *regular pair*[3] for (H, Q, M), the following are equivalent:

a) $y_L \in \arg\min_{y \in Q} \sup_{\nu \in H} M(\nu, y)$,

b) (ν_L, y_L) is a saddle point solution for (H, Q, M).

This theorem establishes that if we exhibit a regular pair whose second argument satisfies a), the game (H, Q, M) has a saddle point, which means that the sequence of max and min in (3.18) can be interchanged. In the following, we find a suitable regular pair, thereby proving (3.19).

Clearly the convexity conditions are satisfied (the set of detectors is not required to be topologized). We need to find a candidate regular pair. Note that the value of inf term in (3.18) is $-\infty$ (which gives a near-far resistance of zero) unless ν is picked such that $\mathcal{R}\nu = u^{k,i}$ (η is invariant with respect to scaling of ν). $u^{k,i}$ is the (k, i)th unit vector in the Hilbert space L, defined as $u_j^{k,i}(l) = \delta_{kj}\delta_{il}$. This gives us a candidate for an optimal detector $\nu_L: d$, with $\mathcal{R}d = u^{k,i}$. Existence of such a vector is shown in (4.6) to follow from the LIA of (2.3).

[3] $(\nu_L, y_L) \in H \times Q$ is a regular pair for (H, Q, M) if, for every $y \in Q$ such that $y_\alpha = (1-\alpha)y_L + \alpha y \in Q$ for $\alpha \in [0, 1]$, we have

$$\sup_{\nu \in H} M(\nu, y_\alpha) - M(\nu_L, y_\alpha) = o(\alpha).$$

(If this detector is indeed optimal, which follows if the candidate pair is regular and satisfies a), and coincides with v^*).

Next we find a y_L which meets the requirement of point a) of the theorem. Using the Cauchy-Schwarz inequality, we find that

$$\sup_{v \in H} M(v, y) = \sup_{\substack{v \in L \\ v^T \mathcal{R} v \neq 0}} \frac{y^T \mathcal{R} v}{\sqrt{v^T \mathcal{R} v}} = \sqrt{y^T \mathcal{R} y} \tag{A.1}$$

where the inner product is maximized for $v = ky + \{x \in L: \mathcal{R}x = 0\}$.

We now need to solve the Hilbert space optimization problem

$$\inf y^T \mathcal{R} y \tag{A.2}$$

$$\text{subject to } y_k(i) = 1.$$

Using (2.21) and the definition of d we can rewrite the minimization problem under consideration as

$$\inf \|y\|_R \tag{A.3}$$

$$\text{subject to } \langle d, y \rangle_R = 1.$$

$\|\cdot\|_R$ is a norm since $\mathcal{R}$ is positive definite. We have obtained a minimum-norm optimization problem in Hilbert space. To prove existence of a solution we need to show that constraint set to be closed, which holds since the Hilbert space is finite dimensional. (Even for $N \to \infty$, when we have an infinite dimensional optimization problem, we could use the fact that the codimension is finite. The problem there is that the signals are no longer square integrable.) The constraint, $y_k(i) = 1$, is equivalent to $y = u^{k,i} + \{x: \langle x, d \rangle_R = 0\}$. $\mathcal{A} = [d]$, the subspace generated by d, is a closed subspace of dimension 1. Hence the constraint set $\{x: \langle x, d \rangle_R = 0\} = \mathcal{A}^\perp$ is closed. We now have a minimum-norm optimization problem in Hilbert space over a closed subspace. Hence, the Projection Theorem, [6], guarantees existence [so we can replace the inf by a min, as required in a)] and uniqueness of a minimizing equivalence class y^*, with

$$y^* \in \{\mathcal{A}^\perp + u^{k,i}\} \cap \mathcal{A}^{\perp\perp} = \{\mathcal{A}^\perp + u^{k,i}\} \cap \mathcal{A} \tag{A.4}$$

where equality holds since A is closed. Hence $y^*(i) = 1$ and $y^* = kd$, which implies

$$y^* = \frac{1}{d_k(i)} d. \tag{A.5}$$

We now have a candidate regular pair which satisfies a): $(v_L, y_L) = (d, (d_k(i))^{-1} d)$. From (A.1) and the definition of regularity we have to check the dependence on α of

$$\begin{aligned}
\sqrt{y_\alpha^T \mathcal{R} y_\alpha} &- \frac{y_\alpha^T \mathcal{R} v_L}{\sqrt{v_L^T \mathcal{R} v_L}} \\
&= \sqrt{d^T \mathcal{R} d + 2\alpha (y - d)^T \mathcal{R} d + \alpha^2 (y - d)^T \mathcal{R} (y - d)} \\
&\quad - \sqrt{\frac{1}{d_k(i)}} \\
&= \sqrt{\frac{1}{d_k(i)} + \alpha^2 (y - d)^T \mathcal{R} (y - d)} - \sqrt{\frac{1}{d_k(i)}}.
\end{aligned} \tag{A.6}$$

We have repeatedly used the decorrelating property of d. Since $\sqrt{1 + x} \leq 1 + 1/2x$, the above quantity lies in the interval $[0, (y - d)^T \mathcal{R} (y - d) \sqrt{d_k(i)}/2\alpha]$, hence divided by α goes to 0 when $\alpha \downarrow 0$. Thus $(d, (d_k(i))^{-1} d)$ is a regular pair which satisfies point a) of the theorem. Hence it follows from the theorem that the penalty function $y^T \mathcal{R} v / \sqrt{v^T \mathcal{R} v}$ has a saddle point, i.e.,

$$\sup_{\substack{v \in L \\ v^T \mathcal{R} v \neq 1}} \inf_{\substack{y \in L \\ y_k(i) = 1}} \frac{y^T \mathcal{R} v}{\sqrt{v^T \mathcal{R} v}} = \inf_{\substack{y \in L \\ y_k(i) = 1}} \inf_{\substack{y \in L \\ v^T \mathcal{R} v \neq 1}} \frac{y^T \mathcal{R} v}{\sqrt{v^T \mathcal{R} v}}. \tag{A.7}$$

◇

Appendix 2

Sufficient Conditions for Linear Independence

Suppose that, for a fixed signal set,

i) $\{\tau_1, \cdots, \tau_K\}$ are continuous random variables,

ii) $\{\tau_1, \cdots, \tau_K\}$ are independent random variables,

iii) $w_k(i) \neq 0$.

Then almost surely there is no $v \in L, v_k(i) \neq 0$ such that $\tilde{S}(t, v) = 0$.

Proof: Define the times of effective arrival and departure of the ith signal of the kth user [1], as

$$\lambda_{i,k}^a = \tau_k + iT + \sup \left\{ \tau \in [0, T), \int_0^\tau s_k^2(t)\, dt = 0 \right\} \tag{A.8}$$

and

$$\lambda_{i,k}^d = \tau_k + iT + \inf \left\{ \tau \in (0, T], \int_\tau^T s_k^2(t)\, dt = 0 \right\}, \tag{A.9}$$

respectively.

Since $v_k(i) \neq 0$ there is a first and a last symbol that differs from zero. It is readily apparent that in order to have $\tilde{S}(t, v) = 0$, the effective arrival of the first (and the effective departure of the last) symbol that differs from zero must be a point of effective multiarrival (respectively multideparture). Note that this property does not depend on the particular v chosen, but only on the set of delays. From (A.8), (A.9), the effective times of arrival and departure inherit from the delays the properties of being continuously valued and mutually independent. Therefore, the result follows, since the set of delays $\{\tau_1, \cdots, \tau_K\}$ for which multiarrival points result has measure zero. ◇

References

[1] S. Verdú, "Minimum probability of error for asynchronous Gaussian multiple-access channels," *IEEE Trans. Inform. Theory*, vol. IT-32, pp. 85-96, Jan. 1986.

[2] —, "Optimum multi-user asymptotic efficiency," *IEEE Trans. Commun.*, vol. COM-34, pp. 890-897, Sept. 1986.

[3] —, "Computational complexity of optimum multi-user detection," *Algorithmica*, vol. 4, pp. 303-312, 1989.

[4] S. Verdú and H. V. Poor, "On minimax robustness: A general approach and applications," *IEEE Trans. Information Theory*, vol. IT-30, pp. 328-340, Mar. 1984.

[5] —, "Minimax robust discrete-time matched filters," *IEEE Trans. Commun.*, vol. COM-31, pp. 208-215, Feb. 1983.

[6] D. Luenberger, *Optimization by Vector Space Methods*. New York: Wiley, 1969.

[7] R. Lupas and S. Verdú, "Linear multiuser detectors for synchronous code-division multiple-access channels," *IEEE Trans. Inform. Theory*, vol. IT-34 1988.

[8] S. Verdú, "Minimum probability of error for asynchronous multiple access communication systems," in *Proc. 1983 IEEE Military Commun. Conf.*, Washington, DC, Nov. 1983, pp. 213-219.

[9] H. V. Poor and S. Verdú, "Single-user detectors for multiuser channels," *IEEE Trans. Commun.*, vol. COM-36, pp. 50-60, Jan. 1988.

[10] M. K. Varanasi and B. Aazhang, "Multistage detection in asynchronous code-division multiple access communications," *IEEE Trans. Commun.*, see this issue, pp. 509-519.

[11] C. Rushforth and Z. Xie, "Multiuser signal detection using sequential decoding," *1988 IEEE Military Commun. Conf.*, San Diego, CA, Oct. 1988.

[12] F. D. Garber and M. B. Pursley, "Optimal phases of maximal-length sequences for asynchronous spread-spectrum multiplexing," *Electron. Lett.*, vol. 16, pp. 756-757, Sept. 1980.

[13] M. B. Pursley and H. F. A. Roefs, "Numerical evaluation of correlation parameters for optimal phases of binary shift-register sequences," *IEEE Trans. Commun.*, vol. COM-27, pp. 1597-1604, Oct. 1979.

[14] E. A. Geraniotis and M. B. Pursley, "Error probability for direct-sequence spread-spectrum multiple-access communications—Part II: Approximations," *IEEE Trans. Commun.*, vol. COM-30, pp. 985–995, May 1982.

[15] R. Lupas, "Near-far resistant linear multiuser detection," Ph.D. dissertation, Princeton Univ., Jan. 1989.

★

Ruxandra Lupas (S'86) was born in Bucharest, Romania, on March 13, 1962. She received the Vordiplom degree in electrical engineering from the Technical University of Munich, West Germany, in 1983, and the Ph.D. degree in electrical engineering from Princeton University in 1989.

She was a Research Associate at DFVLR (German Aerospace Research Establishment), Munich, in the Summer of 1987 and she is currently a member of the Technical Staff at Siemens, Munich, Germany. Her research interests are in communication systems and artificial intelligence.

Ms. Lupas is a member of the German National Scholarship Foundation for having been a winner of the German High-School Competition in Mathematics in 1979, 1980, and 1981. She received an IBM Graduate Fellowship in 1986.

★

Sergio Verdú (S'80–M'84–SM'88) received the Ph.D. degree in electrical engineering from the University of Illinois at Urbana-Champaign in 1984.

Upon completion of his doctorate he joined the Faculty of Princeton University, Princeton, NJ, where he is an Associate Professor of Electrical Engineering. His current research interests are in the areas of multiuser communication, information theory, and statistical signal processing.

Dr. Verdú is a recipient of the National University Prize of Spain, the Rheinstein Outstanding Junior Faculty Award of the School of Engineering and Applied Science in Princeton University, and the NSF Presidential Young Investigator Award. He is currently serving as Associate Editor of the IEEE Transactions on Automatic Control and is a member of the Board of Governors of the IEEE Information Theory Society.

Linear Multiuser Detectors for Synchronous Code-Division Multiple-Access Channels

Ruxandra Lupas
Sergio Verdu

Reprinted from
IEEE TRANSACTIONS ON INFORMATION THEORY
Vol. 35, No. 1, January 1989

Linear Multiuser Detectors for Synchronous Code-Division Multiple-Access Channels

RUXANDRA LUPAS, STUDENT MEMBER, IEEE, AND SERGIO VERDÚ, SENIOR MEMBER, IEEE

Abstract —In code-division multiple-access systems, simultaneous multiuser accessing of a common channel is made possible by assigning a signature waveform to each user. Knowledge of these waveforms enables the receiver to demodulate the data streams of each user, upon observation of the sum of the transmitted signals, perturbed by additive noise. Under the assumptions of symbol-synchronous transmissions and white Gaussian noise, we analyze the detection mechanism at the receiver, comparing different detectors by their bit error rate in the low background noise region, and by their worst-case behavior in a near-far environment where the received energies of the users are not necessarily similar. Optimum multiuser detection achieves important performance gains over conventional single-user detection at the expense of computational complexity that grows exponentially with the number of users. It is shown that in the synchronous case the performance achieved by *linear* multiuser detectors (whose complexity per demodulated bit is only linear in the number of users) is similar to that of optimum multiuser detection. Attention is focused on detectors whose linear memoryless transformation is a generalized inverse of the matrix of signature waveform crosscorrelations, and on the optimum linear detector. It is shown that the generalized inverse detectors exhibit the same degree of near-far resistance as the optimum multiuser detector; the optimum linear detector is obtained subsequently, along with sufficient conditions on the signal energies and crosscorrelations to ensure that its performance is equal to that of the optimum multiuser detector.

I. Introduction

CODE-DIVISION multiple-access is a multiplexing technique where several independent users access simultaneously a multipoint-to-point channel by modulating preassigned signature waveforms. These waveforms are known to the receiver, which observes the sum of the modulated signals embedded in additive white Gaussian noise. If the assigned signals were orthogonal, then a bank of decoupled single-user detectors (matched filters followed by thresholds) would achieve optimum demodulation. In practice, however, orthogonal signal constellations are more the exception than the rule because of bandwidth or complexity limitations (the number of potential users can be very large), lack of synchronism, or other design constraints. Therefore the question of interest is how to demodulate the transmitted messages when the assigned signals are not orthogonal. In practice, demodulation strategies have been restricted to single-user detection, thereby placing the whole burden of complexity on the cross correlation properties of the signal constellation. Recently, the optimum multiuser detector for general asynchronous Gaussian channels was derived and analyzed in [1]. The optimum detector significantly outperforms the conventional single-user detector at the expense of a marked increase in computational complexity—it grows exponentially with the number of users.

The purpose of this paper is to investigate new low-complexity multiuser detection strategies that approach the performance of the optimum detector and to gain further insight into the performance of the optimum multiuser detector. Our attention in focused on symbol-synchronous channels, where the symbol epochs of all users coincide at the receiver. Although in practice this assumption rules out the important class of completely asynchronous code-division multiple-access systems, it holds in slotted channels, and its study is a necessary prerequisite for tackling the general asynchronous channel by allowing us to gain some appreciation of the main issues in the simplest possible setting.

The performance measure of interest is the probability of error of each user. In multiuser problems it is often more convenient and intuitively sound to give information concerning the error probability by means of the *efficiency*, or ratio between the *effective* signal-to-noise ratio (SNR) and the actual SNR, where the effective SNR is the one required to achieve the same probability of error in the absence of interfering users, and the actual SNR is the received energy of the user divided by the power spectral density level of the background thermal white Gaussian noise (not including interference from other users). Note that since the single-user error probability is a one-to-one function of the SNR, the efficiency gives the same information as the error probability. Its limit as the background Gaussian noise level goes to zero, the *asymptotic efficiency*, characterizes the underlying performance loss when the dominant impairment is the existence of other users rather than the additive channel noise. Denoting the power spectral density level of the background white noise by σ^2, the kth user asymptotic efficiency of a detector whose kth user error probability and energy are equal to P_k and w_k,

Manuscript received February 27, 1987; revised October 21, 1987. The material in this paper was partially presented at the 25th IEEE Conference on Decision and Control, Athens, Greece, December 1986. This work was supported in part by the U.S. Army Research Office under Contract DAAL03-87-K-0062.

The authors are with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544.

IEEE Log Number 8825697.

0018-9448/89/0100-0123$01.00

respectively, can be written as [1][1]

$$\eta_k = \sup\left\{0 \le r \le 1;\ \lim_{\sigma \to 0} P_k(\sigma)/Q\left(\frac{\sqrt{r w_k}}{\sigma}\right) < +\infty\right\}, \tag{1.1}$$

i.e., the logarithm of the error probability goes to zero with the same slope as the single-user bit error rate with energy $\eta_k w_k$. In this paper we compare the performance of the various multiuser detectors by means of the asymptotic efficiency. In the high[2] SNR region, the advantage of this measure over the probability of error is twofold. it quantifies the performance degradation due to the existence of other users in a simple, intuitive way, and in contrast to multiuser error probability for which only (asymptotically tight) upper and lower bounds are known [1], exact expressions for the asymptotic efficiency are feasible.

The main shortcoming of currently operational networks employing code-division multiple-access is the *near–far* problem. This refers to the situation wherein the received powers of the users are dissimilar (e.g., in mobile radio networks). Since the output of the matched filter of each user contains a spurious component which is linear in the amplitude of each of the interfering users, the error probability increases to 1/2 as the multiuser interference grows, the asymptotic efficiency becomes zero, and the conventional single-user detector is unable to recover reliably the messages transmitted by the weaker users even if signals with very low crosscorrelations are assigned to the users. However, the near–far problem is not an inherent characteristic of code-division multiple-access systems. Rather, it is the inability of the conventional single-user receiver to exploit the structure of the multiple-access interference that accounts for the ubiquity of the near–far problem in practice. We show that the optimum multiuser detector and other multiuser detectors with much lower computational complexity are *near–far resistant* under mild conditions on the signal constellation. By near–far resistance we mean the asymptotic efficiency minimized over the energies of all the interfering users. If this minimum is nonzero, and, as a consequence, the performance level is guaranteed no matter how powerful the multiuser interference, then we say that the detector is near–far resistant.

The organization of the rest of the paper is as follows. The asymptotic efficiency and the near–far resistance of both the conventional and the optimum detectors are given in Section II. In Section III, we introduce the *decorrelating* multiuser detector. This detector linearly transforms each vector of matched filter outputs with a generalized inverse of the signal crosscorrelation matrix. It is shown that, somewhat unexpectedly, the near–far resistance of the optimum multiuser detector coincides with that of the decorrelating detector whose complexity per demodulated bit is only linear in the number of users. Finally, Section IV investigates the performance of the optimum linear transformation and gives sufficient conditions on the signal energies and crosscorrelations to ensure that the asymptotic efficiency of the optimum linear transformation is equal to that of the optimum multiuser detector.

II. Single-User Detection and Optimum Multiuser Detection

Suppose that the kth user is assigned a finite energy signature waveform, $\{s_k(t),\ t \in [0, T]\}$, and that it transmits a string of bits by modulating that waveform antipodally. If the users maintain symbol synchronization and share a white Gaussian multiple-access channel, then the receiver observes

$$r(t) = \sum_{k=1}^{K} b_k(j) s_k(t - jT) + \sigma n(t), \qquad t \in [jT, jT + T] \tag{2.1}$$

where $n(t)$ is a realization of a unit spectral density white Gaussian process and $\{b_k(j) \in \{-1, 1\}\}_j$ is the kth user information sequence. Assuming that all possible information sequences are equally likely, it suffices to restrict attention to a specific symbol interval in (2.1), e.g., $j = 0$.

It is easy to check that the likelihood function depends on the observations only through the outputs of a bank of matched filters:

$$y_k = \int_0^T r(t) s_k(t)\, dt, \qquad k = 1, \cdots, K \tag{2.2}$$

and therefore $y = (y_1, \cdots, y_K)$ are sufficient statistics for demodulating $b = (b_1, \cdots, b_K)$. We investigate ways of processing these sufficient statistics, which according to (2.1) and (2.2) depend on the transmitted bits in the following way:

$$y = Hb + n \tag{2.3}$$

where H is the nonnegative definite matrix of crosscorrelations between the assigned waveforms:

$$H_{ij} = \int_0^T s_i(t) s_j(t)\, dt \tag{2.4}$$

and its diagonal entries are the energies-per-bit, $H_{ii} = w_i > 0$, of each user; and n is a zero-mean Gaussian K-vector with covariance matrix equal to $\sigma^2 H$.

Conventional single-user detection is the simplest way to make decisions based on y_k; demodulation is decoupled and the multiuser interference is ignored, yielding the following decisions for the kth user:

$$\hat{b}_k^c = \operatorname{sgn} y_k.$$

On the other hand, the optimum multiuser detector selects the most likely hypothesis $\hat{b}^* = (\hat{b}_1^*, \cdots, \hat{b}_K^*)$ given the observations, which corresponds to selecting the noise

[1] $Q(x) = \int_x^\infty (1/\sqrt{2\pi}) e^{-v^2/2}\, dv$.

[2] In the numerical results of [1] and [2], the efficiency is indistinguishable from the asymptotic efficiency for SNR's higher than 7 dB.

realization with minimum energy, i.e.,

$$\hat{b}^* \in \arg \min_{b \in \{-1,1\}^K} \int_0^T \left[r(t) - \sum_{k=1}^{K} b_k s_k(t) \right]^2 dt$$
$$= \arg \max_{b \in \{-1,1\}^K} 2y^T b - b^T H b. \tag{2.5}$$

The computational complexities of the single-user detector and the optimum multiuser detector are radically different. While the time-complexity per bit (TCB) of the single-user detector is independent of the number of users, no algorithm that solves (2.5) in polynomial time in K is known. The reason for this is the nondeterministic polynomial (NP)-completeness of optimum multiuser detection (Appendix I).

The performances of the detectors are also quite different. It is straightforward to find the kth user probability of error of the conventional single-user detector:

$$P_k^c = P[y_k > 0 | b_k = -1]$$
$$= \sum_{\substack{b \in \{-1,1\}^K \\ b_k = -1}} P[y_k > 0, b] P[b | b_k = -1]$$
$$= 2^{1-K} \sum_{\substack{b \in \{-1,1\}^K \\ b_k = -1}} Q\left(\frac{w_k - \sum_{i \neq k} b_i H_{ik}}{\sigma \sqrt{w_k}} \right). \tag{2.6}$$

In the low background noise region, the foregoing summation is dominated by the term corresponding to the least favorable bits of the interfering users, i.e., $b_i = \mathrm{sgn}(H_{ik})$ Thus the asymptotic efficiency of the conventional detector is equal to

$$\eta_k^c = \sup\left\{ 0 \le r \le 1;\ \lim_{\sigma \to 0} P_k^c / Q\left(\frac{\sqrt{r w_k}}{\sigma} \right) < +\infty \right\}$$
$$= \max^2 \left\{ 0, 1 - \sum_{i \neq k} \frac{|H_{ik}|}{w_k} \right\}$$
$$= \max^2 \left\{ 0, 1 - \sum_{i \neq k} |R_{ik}| \frac{\sqrt{w_i}}{\sqrt{w_k}} \right\} \tag{2.7}$$

where R is the matrix of normalized (unit-energy) cross correlations, i.e.,

$$H = W^{1/2} R W^{1/2} \tag{2.8}$$

where $W = \mathrm{diag}\{w_1, \cdots, w_K\}$. It follows from (2.7) that the conventional kth user detector is near-far resistant (i.e., its asymptotic efficiency is bounded away from zero as a function of the interfering users' energies) only if $R_{ik} = 0$ for all $i \neq k$, i.e., only if the kth user's signal is orthogonal to the subspace spanned by the other signals. Otherwise,

$$\bar{\eta}_k^c = \inf_{\substack{w_i \ge 0 \\ i \neq k}} \eta_k^c = 0. \tag{2.9}$$

The kth user error probability of the optimum multiuser receiver is asymptotically (as $\sigma \to 0$) equivalent to that of a binary test between the two closest hypotheses that differ in the kth bit (see [1]). The square of the Euclidean distance between the signals corresponding to these two hypotheses is equal to

$$\min_{b \in \{-1,1\}^K} \min_{\substack{d \in \{-1,1\}^K \\ d_k \neq b_k}} \left| \sum_{i=1}^{K} b_i s_i(t) - \sum_{i=1}^{K} d_i s_i(t) \right|^2$$
$$= 2 \min_{\substack{\epsilon \in \{-1,0,1\}^K \\ \epsilon_k = 1}} \epsilon^T H \epsilon. \tag{2.10}$$

Hence the asymptotic efficiency of the optimum multiuser detector is equal to

$$\eta_k = \frac{1}{w_k} \min_{\substack{\epsilon \in \{-1,0,1\}^K \\ \epsilon_k = 1}} \epsilon^T H \epsilon. \tag{2.11}$$

This is the highest efficiency attainable by any detector because as $\sigma \to 0$ the optimum multiuser detector achieves minimum probability of error for each user. In the two-user case, denoting $\rho = R_{12}$, (2.11) reduces to

$$\eta_1 = \min\left\{ 1, 1 + \frac{w_2}{w_1} - 2|\rho| \frac{\sqrt{w_2}}{\sqrt{w_1}} \right\}, \tag{2.12}$$

and similarly for user 2. Unfortunately, no explicit expressions are known for (2.11) in general. In fact, the combinatorial optimization problem in (2.11) is also NP-complete (Appendix I).

Nevertheless, it is indeed possible to obtain a closed-form expression for the near–far resistance of the optimum multiuser detector, because the minimization of the asymptotic efficiencies with respect to the energies of the interfering user waveforms reduces the combinatorial optimization problem in (2.11) to a continuous optimization problem whose solution is given by the following result.

Proposition 1: Denote by R^+ the Moore–Penrose generalized inverse[3] of the normalized crosscorrelation matrix R. If the signal of the kth user is linearly independent, i.e., it does not belong to the subspace spanned by the other signals, then

$$\bar{\eta}_k = \inf_{\substack{w_i \ge 0 \\ i \neq k}} \eta_k = \frac{1}{R_{kk}^+} \tag{2.13}$$

Otherwise, $\bar{\eta}_k = 0$.

Proof: Using (2.11) for the maximum asymptotic efficiency of the kth user, we obtain

$$\bar{\eta}_k = \min_{\substack{w_i \ge 0 \\ i \neq k}} \min_{\substack{\epsilon \in \{-1,0,1\}^K \\ \epsilon_k = 1}} \frac{1}{w_k} \epsilon^T H \epsilon$$
$$= \min_{\substack{w_i \ge 0 \\ i \neq k}} \min_{\substack{\epsilon \in \{-1,0,1\}^K \\ \epsilon_k = 1}} \frac{1}{w_k} \epsilon^T W^{1/2} R W^{1/2} \epsilon$$
$$= \min_{\substack{x \in R^K \\ x_k = 1}} x^T R x$$
$$= \min_{z \in R^{K-1}} \left(1 + 2 z^T a_k + z^T R_k z \right) \tag{2.14}$$

[3]A generalized inverse A of a matrix B is any matrix that satisfies 1. $ABA = A$ and 2. $BAB = B$. The Moore–Penrose generalized inverse is the unique generalized inverse that satisfies 3. AB and BA are Hermitian.

where R_k is obtained from R by deleting the kth row and column and a_k is the kth column of R with the kth entry removed. Henceforth, we denote such a partitioning of a symmetric matrix with respect to the kth row and column by $R=[R_k, a_k, 1]$, where the rightmost element in the square brackets is the kth diagonal entry. The minimum in the right side of (2.14) is achieved by any element z^* such that

$$R_k z^* = -a_k. \tag{2.15}$$

Because of the Fredholm theorem [11, p. 115], the solvability of (2.15) is equivalent to a_k being orthogonal to the null space of R_k. However, for all $z \in R^{K-1}$ the parabola $q(v) = v^2 + 2vz^T a_k + z^T R_k z$ has at most one zero because it is equal to the quadratic form of the nonnegative definite matrix R with a vector whose kth coordinate is v and whose other components are equal to z. Therefore, the discriminant of the parabola satisfies $(z^T a_k)^2 - z^T R_k z \le 0$; in particular, if z belongs to the null space of R_k, then $z^T a_k = 0$. So a_k is indeed orthogonal to the null space of R_k. Substituting (2.15) into (2.14) we obtain

$$\begin{aligned}\bar{\eta}_k &= 1 - z^{*T} R_k z^* \\ &= 1 - z^{*T} R_k R_k^+ R_k z^* \\ &= 1 - a_k^T R_k^+ a_k. \end{aligned} \tag{2.16}$$

Notice that the kth user is linearly dependent if and only if there exists a linear combination of the columns of R that includes the kth column and is equal to the zero vector. Therefore, if a user is linearly dependent then we can find x such that $Rx = 0$ and $x_k = 1$, in which case the penultimate equation in (2.14) indicates that $\bar{\eta}_k = 0$. To obtain the near–far resistance of a linearly independent user, we employ the following property, which is invoked again later on.

Lemma 1: If the kth user is linearly independent, then every generalized inverse R^I of R satisfies: $(R^I R)_{kj} = \delta_{kj}$, $(RR^I)_{jk} = \delta_{jk}$ and $R_{kk}^I = R_{kk}^+$. (Analogous formulas hold for the unnormalized crosscorrelation matrix H.)

Proof of Lemma 1: Let $S = R^I R - I$. By the definition of generalized inverse, it follows that $RS = 0$, i.e., every column of S is in the null space of R. However, if the kth user is linearly independent, it is necessary that the kth element of each such column be zero. Hence $(R^I R - I)_{kj} = 0$ for all $j = 1, \cdots, K$.

Similarly, with $S = RR^I - I$ and $SR = 0$, we obtain $(RR^I)_{jk} = \delta_{jk}$. Equivalently, $RR^I u_k = u_k$, using the kth unit vector u_k. Hence, for any generalized inverses $R_1^I, R_2^I, R(R_1^I - R_2^I)u_k = 0$. However, since the kth user is linearly independent, it is necessary that the kth element of each vector in the null space of R be zero. Hence $(R_1^I - R_2^I)_{kk} = 0$.

Now we continue with the proof of Proposition 1. Partitioning R^+ with respect to the kth row and column, we have, say, $R^+ = [C, c, \gamma]$. Now, computing the submatrices of the partitioned matrix $R^+ R$ and using Lemma 1, it follows that

$$R_k c + \gamma a_k = 0 \tag{2.17}$$

and

$$c^T a_k + \gamma = 1. \tag{2.18}$$

Notice that $\gamma \neq 0$, for otherwise c would belong to the null space of R_k and would not be orthogonal to a_k, which, as we saw, is not possible. Finally, substituting (2.17) into (2.16) we obtain

$$\begin{aligned}\bar{\eta}_k &= 1 - \frac{1}{\gamma^2} c^T R_k R_k^+ R_k c \\ &= 1 - \frac{1}{\gamma^2} c^T R_k c \\ &= 1 + \frac{1}{\gamma} c^T a_k \\ &= \frac{1}{\gamma} = \frac{1}{R_{kk}^+} \end{aligned} \tag{2.19}$$

where the second, third, and fourth equations follow from the definition of generalized inverse, (2.17) and (2.18), respectively.

III. The Decorrelating Detector

In the absence of noise, the matched filter output vector is $y = Hx$. Thus if the signal set is linearly independent (i.e., H invertible), the natural strategy to follow in this hypothetical situation is to premultiply y by the inverse crosscorrelation matrix H^{-1}. The detector $\hat{x} = \operatorname{sgn} H^{-1} y$ was analyzed in [8], where its performance was quantified in the presence of noise. In [6] it was erroneously shown (cf. [3]) that this detector is optimum in terms of bit-error rate. Note that the noise components in $H^{-1}y$ are correlated, and therefore $\operatorname{sgn} H^{-1} y$ does not result in optimum decisions. It is interesting to point out that this detector does not require knowledge of the energies of any of the active users. To see this, let $\tilde{y}_k = y_k/\sqrt{w_k}$, i.e., $\tilde{y}_k$ is the result of correlating the received process with the normalized (unit-energy) signal of the kth user. Then

$$\begin{aligned}\operatorname{sgn} H^{-1} y &= \operatorname{sgn} W^{-1/2} R^{-1} W^{-1/2} y \\ &= \operatorname{sgn} W^{-1/2} R^{-1} \tilde{y} \\ &= \operatorname{sgn} R^{-1} \tilde{y}, \end{aligned}$$

and therefore, the same decisions are obtained by multiplying the vector of normalized matched filter outputs by the inverse of the normalized crosscorrelation matrix. Apart from the attractive asymptotic efficiency properties shown below for the decorrelating detector, further justification for its study is provided by the fact that it is the solution to the generalized likelihood ratio test or maximum likelihood detector (e.g., [12, ch. 2], [13, p. 291]) when the energies are not known by the receiver. This approach selects the decisions that maximize the maximum of the likelihood function over the unknown parameters, i.e., (cf. (2.5))

$$\begin{aligned}\hat{b}^g &\in \arg \min_{b \in \{-1,1\}^K} \min_{\substack{w_i > 0\\ i=1,\cdots,K}} \int_0^T \left[r(t) - \sum_{k=1}^K b_k s_k(t) \right]^2 dt \\ &= \arg \min_{b \in \{-1,1\}^K} \min_{\substack{w_i > 0\\ i=1,\cdots,K}} y^T H^{-1} y + b^T H b - 2 b^T y \\ &= \arg \min_{b \in \{-1,1\}^K} \min_{\substack{w_i > 0\\ i=1,\cdots,K}} \\ &\cdot \tilde{y}^T R^{-1} \tilde{y} + b^T W^{1/2} R W^{1/2} b - 2 (W^{1/2} b)^T \tilde{y} \\ &= \operatorname{sgn}\left(\arg \min_{x \in R^K} x^T R x - 2 x^T \tilde{y} \right) = \operatorname{sgn} R^{-1} \tilde{y}.\end{aligned}$$

Since in this paper the signal set is not constrained to be linearly independent, the above detector need not exist. In general, we consider the set $I(H)$ of generalized inverses[2] of the crosscorrelation matrix H and analyze the properties of the detector

$$\hat{x} = \operatorname{sgn} H^I y, \tag{3.1}$$

which we refer to as a *decorrelating* detector.

The kth user asymptotic efficiency achieved by a general linear transformation T can be obtained in a way similar to that of the efficiency of the conventional single-user detector $T = I$ (Section II). The first step is to find the bit error probability of the kth user:

$$\begin{aligned}P_k &= P[\hat{x}_k = 1 | x_k = -1] = P[(THx + Tn)_k > 0 | x_k = -1] \\ &= P\left[(Tn)_k > (TH)_{kk} - \sum_{j \neq k} (TH)_{kj} x_j \right] \\ &= 2^{1-K} \sum_{\substack{x \in \{-1,1\}^K \\ x_k = -1}} P\left[(Tn)_k > (TH)_{kk} - \sum_{j \neq k} (TH)_{kj} x_j \right].\end{aligned} \tag{3.2}$$

Since the random variable $(Tn)_k$ is Gaussian with zero mean and variance equal to $(THT^T)_{kk}\sigma^2$, the sum in (3.2) is dominated as $\sigma \to 0$ by the term

$$2^{1-K} Q\left(\left((TH)_{kk} - \sum_{j \neq k} |(TH)_{kj}| \right) \Big/ \sigma \sqrt{(THT^T)_{kk}} \right). \tag{3.3}$$

Hence, according to definition (1.1), the kth user asymptotic efficiency achieved by the linear mapping T is

$$\eta_k(T) = \max{}^2 \left\{ 0, \frac{1}{\sqrt{w_k}} \frac{(TH)_{kk} - \sum_{j \neq k} |(TH)_{kj}|}{\sqrt{(THT^T)_{kk}}} \right\}. \tag{3.4}$$

Thus the kth user asymptotic efficiency of a decorrelating detector with matrix H^I is given by

$$\eta_k(H^I) = \max{}^2 \left\{ 0, \frac{1}{\sqrt{w_k}} \frac{(H^I H)_{kk} - \sum_{j \neq k} |(H^I H)_{kj}|}{\sqrt{(H^I H H^{I^T})_{kk}}} \right\}. \tag{3.5}$$

Proposition 2: If user k is linearly independent every $H^I \in I(H)$ satisfies

$$\eta_k(H^I) = 1/R^+_{kk}. \tag{3.6}$$

Thus for independent users the asymptotic efficiency of the decorrelating detector is independent of the energy of other users and of the specific generalized inverse selected.

Proof: If user k is linearly independent, we established in Lemma 1 that $(H^I H)_{kj} = \delta_{kj}$. Hence it follows from (3.5) that

$$\eta_k(H^I) = \frac{1}{w_k H^I_{kk}}. \tag{3.7}$$

Using the defining properties of generalized inverses (see footnote 2) it is easy to check that if $A \in I(R)$, then $W^{-1/2} A W^{-1/2} \in I(H)$, and if $B \in I(H)$, then $W^{1/2} B W^{1/2} \in I(R)$. Hence there is an obvious bijection between $I(R)$ and $I(H)$. Note that H^+ need not be the image of R^+ in this bijection. However, the inverse image of H^+, say $R^* \in I(R)$, satisfies

$$w_k H^+_{kk} = w_k (W^{-1/2} R^* W^{-1/2})_{kk} = R^*_{kk}. \tag{3.8}$$

Moreover, since user k is linearly independent, Lemma 1 implies that the denominator of (3.7) is equal to the left side of (3.8) and that the right side of (3.8) is equal to R^+_{kk}. Proposition 2 follows.

In Section IV it is shown that if user k is linearly dependent, then

$$\eta^d_k = \sup_{H^I \in I(H)} \eta_k(H^I) = \sup_{T \in R^{K \times K}} \eta_k(T) = \eta^l_k,$$

i.e., the best decorrelating detector and the best linear detector achieve the same kth user asymptotic efficiency.

Proposition 3: The near-far resistance of the decorrelating detector equals that of the optimum multiuser detector, i.e., for all $H^I \in I(H)$,

$$\inf_{\substack{w_j \geq 0\\ j \neq k}} \eta_k(H^I) = \inf_{\substack{w_j \geq 0\\ j \neq k}} \eta_k \equiv \bar{\eta}_k. \tag{3.9}$$

Proof: If user k is linearly independent, then according to Proposition 1 the near-far resistance of the optimum detector is equal to the asymptotic efficiency of the decorrelating detector (Proposition 2), which is independent of the energy of the other users. If user k is linearly dependent, Proposition 1 states that the near-far resistance of the optimum detector is zero, and hence the same is true for any detector.

The result of Proposition 3 is of special importance in a near-far environment, where the received signals have different energies and where the energy ratios may vary continuously over a broad scale if the positions of the users evolve dynamically. In this environment any decorrelating detector, with its linear time-complexity per bit, offers the same near-far resistance as the optimum multiuser detector, whose time-complexity per bit is exponential.

For the case where the signal set is independent, i.e., H is nonsingular (and $\eta_k^d = \eta_k(H^{-1})$ is energy-independent for all users), a geometric explanation for the equality of $\bar{\eta}_k$ and η_k^d can be given in the two-user case. Recall that the received signal y satisfies: $y = Hx + n$ and the noise autocovariance matrix is H. To have spherically symmetric noise, it is convenient to work in the $H^{-1/2}y$ domain. Here the hypotheses, denoted by A, B, C, D in Fig. 1, are at the points $H^{1/2}x$, with $x \in \{-1,1\}^2$. Since in this domain the matched filter output noise is spherically symmetric and Gaussian, the decision regions of the maximum likelihood detector, determined by the minimum Euclidean distance rule, are given by the perpendicular bisectors of the segments between the different hypotheses, and the kth user asymptotic efficiency corresponds to the square of half the minimum distance between distinct hypotheses differing in the kth bit.

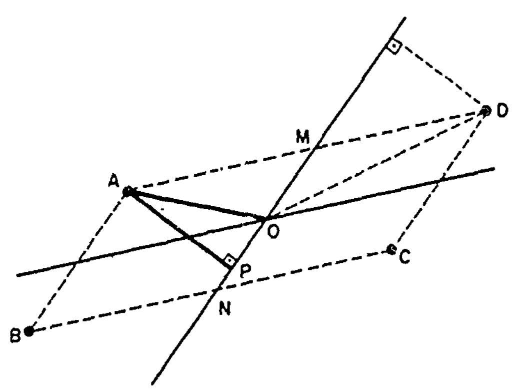


Fig. 1. Hypotheses and decision regions in two-user case.

The decision regions of the decorrelating detector are cones with a vertex at the origin, such that application of H^{-1} maps them to the coordinate axes. Thus in the $H^{-1/2}y$-domain the decision cones pass through the points $H^{1/2}e$, with e the unit vectors in R^2. These points are at the center of the sides of the parallelogram formed by the hypotheses, because the unit vectors can be represented as half the sum of adjacent hypotheses. So, the decorrelating detector decision boundaries are parallel to the parallelogram sides and intersect it at the centers of its sides. The kth bit-error probability (by symmetry we can assume that the transmitted bit was -1) is the sum of two integrals, one for each possibility for the remaining bit, of the noise density function over the region in which the kth bit is decoded as 1. In this case the kth bit-error probability can be easily computed by taking advantage of the aforementioned properties. To this end we rotate the coordinate system to let the y axis coincide with the kth-bit decision boundary and use the equal distance property of the decision boundary to the hypotheses, to observe that the two integrals are equal. We then use the spherical symmetry of the noise to identify each integral as a Q-function of the distance of the hypothesis to the decision boundary. Hence the kth user asymptotic efficiency of the decorrelating detector is equal to the square of the distance of any hypothesis to the kth bit decision boundary. Thus, in Fig. 1, $\sqrt{\eta_1}$ is the length of the shortest of the segments AM, AO and BO, and $\sqrt{\eta_1^d}$ is the length of AP. The result of Proposition 3 can now be interpreted as follows. Since η appears as the hypotenuse and η^d as the leg of a right-angled triangle, η is lower-bounded by the energy independent η^d. However, since the triangle angles vary with increasing energy of the interfering user waveform, there is a particular energy ratio for which the triangle degenerates into a line segment. This is the point when η reaches its minimum $\bar{\eta}$, which is geometrically identical with η^d. For the parallelogram formed by the hypotheses, this is the case where a diagonal is perpendicular to a side (e.g., AO perpendicular to CD).

IV. The Optimum Linear Multiuser Detector

We now turn to the question of finding the optimum linear detector. We have seen that this is a fruitful approach, since a particular type of linear detector, the decorrelating detector, offered a substantial improvement in asymptotic efficiency compared to the single-user detector, while its near–far resistance equaled that of the optimum multiuser detector. While we now know that no detector, linear or nonlinear, can outperform the decorrelating detector with respect to near–far resistance, for fixed energies it is indeed possible to obtain linear detectors that have a higher asymptotic efficiency than the one achieved by the decorrelating detector.

We find the linear detector which maximizes the asymptotic efficiency (or equivalently minimizes the probability of bit error in the low-noise region) and compare the achieved asymptotic efficiency to the ones achieved by the conventional and optimal detectors. Thus we ask which mapping T: $R^K \to R^K$ maximizes the asymptotic efficiency of the decision scheme

$$\hat{x} = \mathrm{sgn}(Ty) = \mathrm{sgn}(THx + Tn). \tag{4.1}$$

The interpretation of this optimization problem in terms of decision regions is to find the optimal partition of the K-dimension hypotheses space into K decision cones with vertices at the origin. The surfaces of these cones determine the columns of the inverse T^{-1} of the mapping sought. Application of T on the cone configuration will map the cones on quadrants, after which a sign detector is used.

The kth user asymptotic efficiency of a general linear detector, as given by (4.1) was derived in (3.4):

$$\eta_k(T) = \max{}^2\left\{0, \frac{1}{\sqrt{w_k}} \frac{(TH)_{kk} - \sum_{j\neq k} |(TH)_{kj}|}{\sqrt{(THT^T)_{kk}}}\right\}. \tag{4.2}$$

The best linear detector has the asymptotic efficiency

$$\eta_k^l = \sup_{T \in R^{K\times K}} \eta_k(T). \tag{4.3}$$

Hence the asymptotic efficiency of the best linear detector is equal to

$$\eta_k^l = \sup_{v \in R^K} \max{}^2 \left\{ 0, \frac{1}{\sqrt{w_k}} \frac{h_k^T v - \sum_{j \neq k} |h_j^T v|}{\sqrt{v^T H v}} \right\}$$
$$= \max{}^2 \left\{ 0; \sup_{v \in R^K} \eta_k(v) \right\} \quad (4.4)$$

with

$$\eta_k(v) = \frac{1}{\sqrt{w_k}} \frac{h_k^T v - \sum_{j \neq k} |h_j^T v|}{\sqrt{v^T H v}} \quad (4.5)$$

where v denotes the kth row of T. To minimize the probability of P_k, we have to maximize the argument of the Q-function, and equivalently maximize the asymptotic efficiency $\eta_k(v)$, with respect to the components of the vector v. Since the map applied on the matched filter outputs is linear, the asymptotic efficiencies of all the users can be simultaneously maximized, each such maximization yielding the corresponding row of the map to be applied. For the sake of clarity, we first consider the two-user case, for which explicit expressions for the maximum linear asymptotic efficiency can be obtained.

A. The Two-User Case

Throughout this subsection we denote the normalized crosscorrelation between the signals by $\rho = R_{12}$. We first give an explicit expression for the optimum linear detector.

Proposition 4: The kth user optimal linear transformation $T_k(y) = v^T y$ on the matched filter outputs prior to threshold detection is given by

$$v^T = \left[1; -\operatorname{sgn}\rho \min\left\{1, |\rho|(w_k/w_i)^{1/2}\right\}\right] \quad (4.6)$$
$$= \begin{cases} [1; -\operatorname{sgn}\rho], & \text{if } (w_i/w_k)^{1/2} \le |\rho| \\ b_k^T, & \text{otherwise} \end{cases} \quad (4.7)$$

where b_k^T is the kth row of the decorrelating detector and $(i,k) \in \{(1,2),(2,1)\}$.

Proof. Without loss of generality, let $k = 1$. We have

$$H = \begin{bmatrix} w_1 & \rho\sqrt{w_1 w_2} \\ \rho\sqrt{w_1 w_2} & w_2 \end{bmatrix}, \qquad v^T = [1; v_2] \quad (4.8)$$
$$\eta_1(v) = \frac{1}{\sqrt{w_1}} \frac{h_1^T v - |h_2^T v|}{\sqrt{v^T H v}}$$
$$= \frac{1 + \rho(w_2/w_1)^{1/2} v_2 - |\rho(w_2/w_1)^{1/2} + (w_2/w_1) v_2|}{\sqrt{1 + 2\rho(w_2/w_1)^{1/2} v_2 + (w_2/w_1) v_2^2}} \quad (4.9)$$

and the objective is to maximize the right side of (4.9) with respect to v_2. We consider the case $|\rho| = 1$ separately.

a) Case $|\rho| \neq 1$*:* Introduce an indicator function for the absolute value term:

$$I = \begin{cases} 1, & \rho + (w_2/w_1)^{1/2} v_2 > 0 \\ -1, & \rho + (w_2/w_1)^{1/2} v_2 < 0 \\ 0, & \text{else} \end{cases}. \quad (4.10)$$

Then

$$\frac{d\eta_1}{dv_2} = -\frac{(1-\rho^2)(w_2/w_1)}{\left(1 + 2\rho(w_2/w_1)^{1/2} v_2 + (w_2/w_1) v_2^2\right)^{3/2}} (I + v_2). \quad (4.11)$$

Therefore, we should take $v_2 = -I$ when this is consistent with the definition of I as a function of v_2. Thus

$$v_2 = \begin{cases} 1, & \text{if } I = -1 \Leftrightarrow 0 < (w_2/w_1)^{1/2} < -\rho \\ -1, & \text{if } I = 1 \Leftrightarrow 0 < (w_2/w_1)^{1/2} < \rho \end{cases}. \quad (4.12)$$

As can easily be seen, both values correspond to maxima. If neither of these conditions is met, the derivative does not have a zero. The optimal value for v_2 can be determined by taking a closer look at the behavior of $d\eta_1/dv_2$, in Fig. 2.

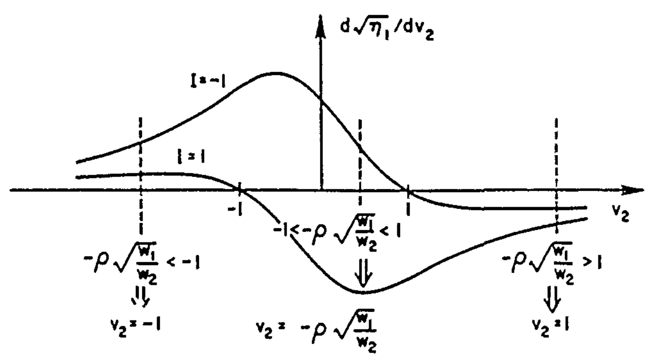


Fig. 2. Behavior of derivative in (4.11).

For both $I = 1$ and $I = -1$, the derivative of η_1 is positive for v_2 smaller than the abscissa of the zero of the derivative (which is equal to $-I$), and negative afterwards. Due to the nonlinearity of η_1 the derivative has the form corresponding to $I = -1$ for $v_2 < -\rho(w_2/w_1)^{1/2}$ and the form corresponding to $I = 1$ afterwards. Since the second branch (for $I = 1$) turns negative before the first one, we have to take the largest value of v_2 yielding a positive derivative on the first branch. It can easily be seen that in the "no-zero" case, $-1 < -\rho(w_2/w_1)^{1/2} < 1$, this is the point of discontinuity, i.e., $v_2 = -\rho(w_2/w_1)^{1/2}$. Note that for $\rho = 0$ we get $v^T = [1; 0]$, the identity transformation, as expected, since the users are then decoupled and a single-user detector is optimal. By taking the inverse of R we also see that in the no-zero case the optimal transformation vector is exactly the corresponding row of the inverse correlation matrix.

b) Case $|\rho| = 1$: Equation (4.9) becomes

$$\eta_1(v_2) = \mathrm{sgn}\left(1 + \mathrm{sgn}\,\rho (w_2/w_1)^{1/2} v_2\right) - (w_2/w_1)^{1/2}. \tag{4.13}$$

We see that, for $(w_2/w_1)^{1/2} < 1$, any v_2 satisfying v_2 sgn $\rho > -(w_1/w_2)^{1/2}$ is optimal, in particular the one given in (4.7). Otherwise, the asymptotic efficiency of the best linear transformation is zero, hence all linear transformations are equivalent.

Substituting the result of Proposition 4 into the asymptotic efficiency of (4.9), we obtain the following

Proposition 5: The kth user asymptotic efficiency of the optimal linear two-user detector equals

$$\eta_k^l = \begin{cases} 1 - 2|\rho|(w_i/w_k)^{1/2} + w_i/w_k, & \text{if } (w_i/w_k)^{1/2} \le |\rho| \\ 1-\rho^2, & \text{otherwise} \end{cases} \tag{4.14}$$

for $(i,k) \in \{(1,2),(2,1)\}$.

The kth user asymptotic efficiency obtained in the range $(w_i/w_k)^{1/2} < |\rho|$ equals the optimum asymptotic efficiency, obtained in (2.12). Even outside the region of optimality, the best linear detector shows a far better performance than the conventional single-user detector (see Fig. 3), since if $w_i/w_k > \rho^2$, then η_1^l is independent of w_i/w_k, whereas according to (2.7) the asymptotic efficiency of the conventional detector is equal to zero for $w_i/w_k \ge 1/\rho^2$.

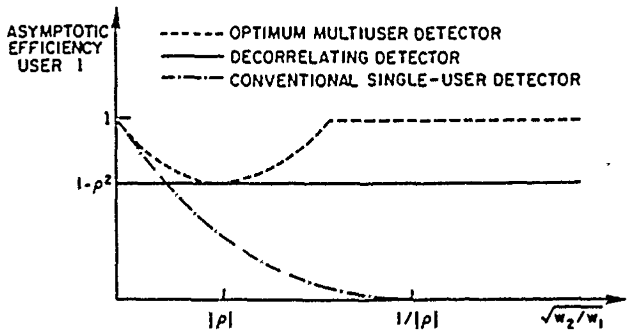


Fig. 3. Asymptotic efficiencies in two-user case ($\rho = 0.6$).

There is an intuitive interpretation of the dual behavior of the best linear detector and of the boundary point $(w_i/w_k)^{1/2} = |\rho|$. Let $k=1$. The input to the threshold device corresponding to the first user, $z_1 = v^T y$, has three components:

$$z_1 = w_1\left[(1-\rho^2) + \rho\left(\rho + v_2 (w_2/w_1)^{1/2}\right)\right] x_1 + w_1\left[(w_2/w_1)^{1/2}\left(\rho + v_2(w_2/w_1)^{1/2}\right)\right] x_2 + \tilde{n} \tag{4.15}$$

where $\tilde{n}$ is a Gaussian random variable of variance $w_1\sigma^2[(1-\rho^2) + (\rho + v_2(w_2/w_1)^{1/2})^2]$. For $(w_2/w_1)^{1/2} > |\rho|$, the second term outweighs the second part of the first term, so the best one can do is to eliminate it by choosing $v_2 = -\rho(w_1/w_2)^{1/2}$ (the decorrelating detector). Since this minimizes the noise variance at the same time, it is the best strategy in this region. If, however, $(w_2/w_1)^{1/2} < |\rho|$, and v_2 is such that the term $\rho(\rho + v_2(w_2/w_1)^{1/2})$ is positive, it is a better policy to allow interference from user 2, which is compensated by the second part in the first term, and use the residual positive contribution in the first term to increase the SNR as compared to the decorrelating case. We have seen that this strategy leads to the same performance as the more complex maximum likelihood detector.

Note that in the two-user case the signal energies and cross correlations cannot be picked so as to allow *both* users optimal performance at the same time, for user 1 we need $(w_2/w_1)^{1/2} < |\rho| < 1$, whereas for user 2 we need $(w_2/w_1)^{1/2} > 1/|\rho| > 1$.

B. The K-User Case

Unlike Propositions 2 and 5, in the general K-user case it is not feasible to obtain an explicit expression for the asymptotic efficiency achieved by the best linear detector

Proposition 6. The kth user asymptotic efficiency of the best linear detector equals:

$$\eta_k^l = \frac{1}{w_k} \max{}^2\left\{0, \max_{\substack{e_j \in \{-1,1\} \\ j \ne k}} \eta(e)\right\} \tag{4.16a}$$

with

$$\eta(e) = \max_{\substack{v \in R^K \\ v^T H v = 1 \\ e_j h_j^T v \ge 0 \\ j \ne k}} v_o^T H v \tag{4.16b}$$

where the ith component of v_o is equal to

$$(v_o)_i = \begin{cases} -e_i, & i \ne k \\ 1, & i = k \end{cases}.$$

Then the maximum $\eta(e)$ is achieved for $\tilde{v}$ such that

$$\tilde{v} = \frac{v_o + \sum_{j\ne k} \lambda_j e_j u_j}{\left(v_o^T H v_o + v_o^T H \sum_{j \ne k} \lambda_j e_j u_j\right)^{1/2}},$$

$$(u_j)_i = \begin{cases} 0, & i \ne j \\ 1, & i = j \end{cases} \tag{4.17}$$

$$e_j h_j^T \tilde{v} \ge 0 \text{ for } j \ne k \tag{4.18}$$

$$h_j^T \tilde{v} \ne 0 \Rightarrow \lambda_j = 0 \tag{4.19}$$

$$\lambda_j \ge 0, \qquad j \ne k. \tag{4.20}$$

Proof: Let

$$S_j^+ = \left\{x \in R^K : h_j^T x \ge 0\right\}$$
$$S_j^- = \left\{x \in R^K : h_j^T x \le 0\right\}. \tag{4.21}$$

From (4.5) we seek

$$\sup_{v\in R^K}\frac{1}{\sqrt{v^THv}}\left(h_k^Tv-\sum_{j\neq k}|h_j^Tv|\right)$$

$$=\max_{\substack{e_j\in\{-1,1\}\\ j\neq k}}\ \sup_{v\in\cap_j S_j^{e_j}}\frac{1}{\sqrt{v^THv}}\left(h_k^Tv-\sum_{j\neq k}|h_j^Tv|\right)\quad(4.22)$$

$$=\max_{\substack{e_j\in\{-1,1\}\\ j\neq k}}\eta(e),\ \text{with}\ \eta(e)$$

$$=\sup_{v\in\cap_j S_j^{e_j}}\frac{1}{\sqrt{v^THv}}\left(h_k^Tv-\sum_{j\neq k}|h_j^Tv|\right).\quad(4.23)$$

From the definition of v_o we see that the term in parentheses equals v_0^THv. Now $v\in\cap_j S_j^{e_j}\Leftrightarrow e_jh_j^Tv\geq 0,\ j\neq k$, and since η_k^l is invariant to scaling of v, maximization of the given functional over R^K is equivalent to maximization over the ellipsoid $v^THv=1$.

This proves the first part of Proposition 6. We now have to perform two maximizations where the second one has the explicit form of an exhaustive search. We turn our attention to the inner maximization in (4.16). We first show that it is possible to replace the feasible set therein by an equivalent convex set, i.e., the asymptotic efficiency is unchanged if we replace

$$\eta(e)=\sup_{\substack{v\in R^K\\ v^THv=1\\ e_jh_j^Tv\geq 0\\ j\neq k}}v_0^THv\ \text{by}\ \eta(e)=\sup_{\substack{v\in R^K\\ v^THv\leq 1\\ e_jh_j^Tv\geq 0\\ j\neq k}}v_0^THv.\quad(4.24)$$

To show (4.24), let $y=H^{1/2}v$, $z_j^T=j$th row of $H^{1/2}$. It then follows that $h_j^Tv=z_j^Ty$, $v_0^TH^{1/2}=y_o^T$, $v^THv=y^Ty=|y|^2$, and

$$\eta(e)=\sup_{\substack{y\in R^K\\ |y|=1\\ e_jz_j^Ty\geq 0\\ j\neq k}}y_o^Ty=\sup_{\substack{y\in R^K\\ |y|=1\\ e_jz_j^Ty\geq 0\\ j\neq k}}|y_o||y|\cos\alpha\quad(4.25)$$

where α is the angle between the vectors y_o and y. Since the inequality constraints are linear and partition the space into convex cones with vertex at the origin, the optimal angle α is independent of $|y|$. Either the optimal $\cos\alpha$ is nonnegative, in which case $\eta(e)$ is maximized for $|y|$ maximal in both versions, or it is negative, in which case $\eta(e)<0$. In either case, the value of η_k^l, which involves comparison with zero, is unchanged if the maximization is performed over the interior of the ellipsoid, which completes the proof of the claim.

We now have to consider the following problem:

$$\eta(e)=\inf_{\substack{v\in R^K\\ v^THv-1\leq 0\\ -e_jh_j^Tv\leq 0\\ j\neq k}}-v_0^THv.\quad(4.26)$$

Since this is a minimization problem of a continuous convex function on a compact convex set, it achieves a unique minimum on the set. Since all the functions are differentiable, we can apply the Kuhn–Tucker conditions (e.g., [4]), to get from condition (1),

$$-Hv_o+\lambda_o2H\tilde{v}-\sum_{j\neq k}\lambda_je_jh_j=0,$$

hence

$$\tilde{v}=\frac{1}{2\lambda_o}\left(v_o+\sum_{j\neq k}\lambda_je_ju_j\right)\quad(4.27)$$

with u_j the jth unit vector, as defined before. Equations (4.18) and (4.19) result from the Kuhn–Tucker conditions, condition (4.20) expresses the nonnegativity requirement for the λ_i. There is one more constraint to satisfy, which is $\tilde{v}^TH\tilde{v}=1$:

$$1=\tilde{v}^TH\tilde{v}=\frac{1}{2\lambda_o}\left(v_o^TH\tilde{v}+\sum_{j\neq k}\lambda_je_jh_j^T\tilde{v}\right)=\frac{v_o^TH\tilde{v}}{2\lambda_o}.$$

We used condition (4.19) to get the last equality, so

$$2\lambda_o=v_o^TH\tilde{v}=\eta(e),\quad(4.28)$$

and since

$$v_o^TH\tilde{v}=\frac{1}{2\lambda_o}\left(v_o^THv_o+\sum_{j\neq k}\lambda_je_jh_j^Tv_o\right),$$

we get

$$2\lambda_o=\left(v_o^THv_o+v_o^TH\sum_{j\neq k}\lambda_je_ju_j\right)^{1/2}.$$

This together with (4.27) completes the proof of Proposition 6.

In Appendix II we show an explicit procedure for finding the best linear detector characterized in Proposition 6. Its asymptotic efficiency is trivially upper- and lower-bounded by that of the optimum and decorrelating detectors, respectively. For certain values of energies and crosscorrelations these bounds are attained; sufficient conditions for this to occur are given in Propositions 7 and 8.

Proposition 7: The following are sufficient conditions on the signal energies and crosscorrelations for the best linear detector to achieve optimal kth user asymptotic efficiency:

$$\sqrt{w_k}>\max_{j=1,\cdots,K}\left(\frac{1}{|R_{kj}|}\sum_{i\neq k}\sqrt{w_i}|R_{ij}|\right).\quad(4.29)$$

Proof: In the optimality case, we show in Appendix II that $e_jh_j^Tv_o>0$ for all $j\neq k$. If we introduce $e_k=1$ this has to hold also for $j=k$, otherwise we get negative asymptotic efficiency. We can rewrite these conditions as

$$DHD(-1-1\cdots 1\cdots -1)^T = \begin{bmatrix} H_{11} & e_1e_2H_{12} & \cdots & e_1H_{1k} & \cdots & e_1e_KH_{1K} \\ e_1e_2H_{21} & H_{22} & \cdots & e_2H_{2k} & \cdots & e_2e_KH_{2K} \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ e_1e_KH_{K1} & e_Ke_2H_{K2} & \cdots & e_KH_{Kk} & \cdots & H_{KK} \end{bmatrix} \begin{bmatrix} -1 \\ -1 \\ \cdot \\ 1 \\ \cdot \\ -1 \end{bmatrix} > 0 \qquad (4.30)$$

where D is the diagonal matrix with ith diagonal element equal to e_i. We now see that a sufficient condition for the above inequality to hold for some $e_1,\cdots,e_K$ is

$$|H_{jk}| > \sum_{i \neq k} |H_{ji}|, \qquad j=1,\cdots,K.$$

The corresponding e_j are $e_j = \operatorname{sgn} H_{jk}$. Hence (4.29) follows by replacing H_{ij} by $\sqrt{w_iw_j}\,R_{ij}$.

Note that the above condition can be satisfied by only one user because

$$\sqrt{w_k} > \sqrt{w_j}/|R_{kj}| > \sqrt{w_j}, \qquad \text{for all } j.$$

Proposition 8: If user k is linearly independent, the following condition is sufficient for the kth row of the decorrelating detector $H^I \in I(H)$ to be the best kth user linear detector for a given set of signal energies and crosscorrelations:

$$|H^I_{jk}| \le H^I_{kk}, \qquad \text{for all } j \neq k. \qquad (4.31)$$

Proof: We showed that in the terminal case $v = h^I_k/\sqrt{H^I_{kk}}$ is a maximizing vector for v_o^THv, hence there are nonnegative Kuhn–Tucker multipliers λ_k, such that, with (4.27) and (4.28),

$$v = \frac{h^I_k}{\sqrt{H^I_{kk}}} = \sqrt{H^I_{kk}}\left(v_o + \sum_{j \neq k} \lambda_j e_j u_j\right)$$

or

$$\frac{1}{H^I_{kk}} h^I_k = [(\lambda_1 - 1)e_1,\cdots,1,\cdots,(\lambda_K - 1)e_K]^T$$

so

$$\lambda_j = 1 + e_jH^I_{jk}/H^I_{kk}, \qquad j \neq k. \qquad (4.32)$$

Hence (4.31) is sufficient to ensure $\lambda_j \ge 0$ regardless of $\{e_i, i \neq k\}$.

Note that in the two-user case, Proposition 5 implies that the sufficient conditions found in Propositions 7 and 8 are also necessary.

Proposition 9: If user k is linearly dependent, then

$$\eta^d_k = \sup_{H^I \in I(H)} \eta_k(H^I) = \sup_{T \in R^{K\times K}} \eta_k(T) = \eta^l_k \quad (4.33)$$

i.e., for a dependent user the best decorrelating detector has the same asymptotic efficiency as the best linear detector.

Proof: Recall the bijection between $I(R)$ and $I(H)$ established in Section III. Let R^I denote the image of H^I under this bijection. Then, using (3.5) for the last equality, we can write

$$\eta^d_k = \sup_{H^I \in I(H)} \eta_k(H^I) = \sup_{R^I \in I(R)} \eta_k(W^{-1/2}R^IW^{-1/2})$$

$$= \max{}^2\left\{0, \sup_{R^I \in I(R)} \frac{(R^IR)_{kk} - \sum_{j \neq k} |(R^IR)_{kj}|\frac{\sqrt{w_j}}{\sqrt{w_k}}}{\sqrt{(R^IRR^{I^T})_{kk}}}\right\}. \qquad (4.34)$$

Since R is nonnegative definite of rank r, it can be written using its orthonormal eigenvector matrix T and the $r \times r$ diagonal matrix Λ of nonzero eigenvalues of R, as

$$R = T\begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix}T^T. \qquad (4.35)$$

Then (cf. [10]), R^I is a generalized inverse of R if and only if, for some matrices U and V of appropriate dimensions, it can be written as

$$R^I = T\begin{bmatrix} \Lambda^{-1} & V \\ U & U\Lambda V \end{bmatrix}T^T. \qquad (4.36)$$

Hence, using the corresponding partition of T, we can write

$$(R^IR)_{kj} = u_k^T[T_1\ T_2]\begin{bmatrix} I & 0 \\ U\Lambda & 0 \end{bmatrix}\begin{bmatrix} T_1^T \\ T_2^T \end{bmatrix}u_j = u_k^T(T_1T_1^T + T_2U\Lambda T_1^T)u_j \qquad (4.37)$$

$$(R^IRR^{I^T})_{kk} = u_k^T[T_1\ T_2]\begin{bmatrix} \Lambda^{-1} & U^T \\ U & U\Lambda U^T \end{bmatrix}\begin{bmatrix} T_1^T \\ T_2^T \end{bmatrix}u_k = u_k^T(T_1\Lambda^{-1}T_1^T + T_2UT_1^T + T_1UT_2^T + T_2U\Lambda U^TT_2^T)u_k \qquad (4.38)$$

and

$$\eta^d_k = \max{}^2\left\{0, \sup_{U \in R^{(K-r)\times r}} \frac{u_k^T(T_1 + T_2U\Lambda)T_1^Tu_k - \sum_{j \neq k} |u_k^T(T_1 + T_2U\Lambda)T_1^Tu_j|\frac{\sqrt{w_j}}{\sqrt{w_k}}}{\sqrt{u_k^T(T_1 + T_2U\Lambda)\Lambda^{-1}(T_1 + T_2U\Lambda)^Tu_k}}\right\}. \qquad (4.39)$$

Since user k is dependent, $u_k^T T_2$, whose components are the kth components of the eigenvectors to eigenvalue zero, is nonzero. (Otherwise, for all x with $Rx = 0$, x_k would be zero, which implies that the kth user is linearly independent of the other users.) Hence since Λ is invertible, we can make the change of variables

$$x = (T_1 + T_2 U \Lambda)^T u_k \tag{4.40}$$

to get

$$\eta_k^d = \max{}^2 \left\{ 0, \sup_x \frac{x^T T_1^T u_k - \sum_{j \neq k} |x^T T_1^T u_j| \frac{\sqrt{w_j}}{\sqrt{w_k}}}{\sqrt{x^T \Lambda^{-1} x}} \right\}. \tag{4.41}$$

Using the same reasoning as in the proof of Proposition 6 for the best linear detector, we can write

$$\eta_k^d = \max{}^2 \left\{ 0, \max_{\substack{e_j \in \{-1,1\} \\ j \neq k}} \sup_{\substack{x \in R^r \\ x^T \Lambda^{-1} x = 1 \\ e_j x^T T_1^T u_j \geq 0 \\ j \neq k}} \frac{1}{w_k} v_o^T W^{1/2} T_1 x \right\} \tag{4.42}$$

where the ith component of v_o is equal to

$$(v_o)_i = \begin{cases} -e_i, & i \neq k \\ 1, & i = k \end{cases},$$

$$\eta_k^d = \frac{1}{w_k} \max{}^2 \left\{ 0, \max_{\substack{e_j \in \{-1,1\} \\ j \neq k}} \eta_k^d(e) \right\} \tag{4.43a}$$

with

$$\eta_k^d(e) = \sup_{\substack{x \in R^r \\ x^T \Lambda^{-1} x = 1 \\ e_j x^T T_1^T u_j \geq 0 \\ j \neq k}} v_o^T W^{1/2} T_1 x \tag{4.43b}$$

whereas the kth user asymptotic efficiency of the best linear detector equals (cf. (4.16)),

$$\eta_k^l = \frac{1}{w_k} \max{}^2 \left\{ 0, \max_{\substack{e_j \in \{-1,1\} \\ j \neq k}} \eta_k^l(e) \right\}$$

with

$$\eta_k^l(e) = \sup_{\substack{v \in R^K \\ v^T H v = 1 \\ e_j h_j^T v \geq 0 \\ j \neq k}} v_o^T H v.$$

Let

$$v^* \in \arg \eta_k^l(e) = \arg \max_{\substack{v \in R^K \\ v^T H v = 1 \\ e_j h_j v \geq 0 \\ j \neq k}} v_o^T H v. \tag{4.44}$$

We show that $x^* = \Lambda T_1^T W^{1/2} v^*$ is feasible, and that $v_o^T W^{1/2} T_1 x^* = \eta_k^l(e)$:

$$\begin{aligned} e_j x^{*T} T_1^T u_j &= e_j v^{*T} W^{1/2} T_1 \Lambda T_1^T u_j \\ &= e_j v^{*T} W^{1/2} R W^{1/2} W^{-1/2} u_j \\ &= \frac{1}{\sqrt{w_j}} e_j v^{*T} H u_j \\ &= \frac{1}{\sqrt{w_j}} e_j v^{*T} h_j \geq 0 \end{aligned} \tag{4.45}$$

since v^* is feasible,

$$\begin{aligned} x^{*T} \Lambda^{-1} x^* &= v^{*T} W^{1/2} T_1 \Lambda \Lambda^{-1} \Lambda T_1^T W^{1/2} v^* \\ &= v^{*T} H v^* = 1. \end{aligned} \tag{4.46}$$

Hence x^* is feasible, and

$$\begin{aligned} v_o^T W^{1/2} T_1 x^* &= v_o^T W^{1/2} T_1 \Lambda T_1^T W^{1/2} v^* \\ &= v_o^T H v^* = \eta_k^l(e). \end{aligned} \tag{4.47}$$

We know that $\eta_k^d \leq \eta_k^l$, since the decorrelating detector belongs to the class of linear detectors. We exhibited for each e a feasible vector x^*, which satisfied $v_o^T W^{1/2} T_1 x^* = \eta_k^l(e)$. Since from (4.43), $\eta_k^d(e) \geq v_o^T W^{1/2} T_1 x$ for all feasible x, we have, for all e, $\eta_k^d(e) \geq \eta_k^l(e)$. Hence $\eta_k^d \geq \eta_k^l$, which establishes (4.33).

Since the kth user asymptotic efficiency depends only on the kth row of the applied linear transformation, optimization of $\eta_k(H^I)$ over the class of generalized inverses for each dependent user k, yields different rows, each belonging to a different generalized inverse. Consequently, the collection of the K optimal rows need not be a generalized inverse.

Finally, notice that the near–far resistance of the optimum linear detector is equal to that of the optimum detector, since it is shown in Proposition 3 that a particular type of linear detector, namely, the decorrelating detector, achieves optimum near–far resistance.

V. Conclusion

The main contribution of this paper is the establishment of the fact that a set of appropriately chosen memoryless linear transformations on the outputs of a matched filter bank exhibits a substantially higher performance than the conventional single-user detector, while maintaining a comparable ease of computation. Moreover, the near–far resistance of all proposed detectors is shown to equal that of the optimum multiuser detector.

Even though the worst-case complexity of the algorithm used to find the best linear detector is exponential in the number of users, in a fixed-energy environment this computation needs to be carried out only once; hence the real-time time-complexity per bit is linear, in contrast to the optimum multiuser detector. Moreover, a region of signal energies and crosscorrelations exists in which the optimal linear detector achieves optimum asymptotic efficiency.

The decorrelating detector is easier to compute than the optimum linear detector, and it exhibits either the same or quite similar performance, depending on the energies and correlations. Since the decorrelating detector does not require knowledge of the transmitters' energies and it achieves the highest possible degree of near–far resistance, it is an attractive alternative to the optimum detector in situations where the received energies are not fixed. The only requirement for the signal of a user to be detected reliably by the decorrelating detector regardless of the level of multiple-access interference, is that it does not belong to the subspace spanned by the other signals—a mild constraint that should be compared to the condition necessary for reliable detection by the conventional single-user detector, i.e., that the signal is *orthogonal* to all the other signals.

The most interesting generalization of the results of this paper is the *asynchronous* code-division multiple-access channel.[4] Due to the fact that in the asynchronous case the channel has memory, a K-input K-output linear discrete-time filter will replace the memoryless linear transformation studied in this paper.

Appendix I

This appendix gives a summary of the results in [18]. We show that the problems of optimum multiuser demodulation and solving for the maximum asymptotic efficiency are nondeterministic polynomial time hard (NP-hard) in the number of users and therefore do not admit polynomial time algorithms unless such algorithms are found for a large class of well-known combinatorial problems including the traveling salesman and integer linear programming. According to (2.5), the selection of the most likely hypothesis given the observations is the following combinatorial optimization problem.

MULTIUSER DETECTION—

Instance: Given $K \in Z^+$, $y \in Q^K$ and a nonnegative definite matrix $H \in Q^{K \times K}$;

Find $\{b^* \in \{-1,1\}^K\}$ that maximizes $2b^T y - b^T H b$.

Proposition 10. MULTIUSER DETECTION is NP-hard.

Proof: The proof of NP-hardness of MULTIUSER DETECTION can be carried out by direct transformation from the following NP-complete problem [15].

PARTITION—

Instance: Given $L \in Z^+$, $\{l_i \in Z^+, i = 1,\cdots,L\}$;

Question: Is there a subset $I \subset \{1,\cdots,L\}$ such that $\Sigma_{i \in I} l_i = \Sigma_{i \notin I} l_i$?

Given $l_1,\cdots,l_L$, we choose the following instance of MULTIUSER DETECTION:

$$K = L$$
$$h_{ij} = l_i l_j$$
$$y_k = 0, \quad k = 1,\cdots,K.$$

With this choice, $\{l_1,\cdots,l_L\}$ is a "yes" instance of PARTITION if and only if

$$\max_{b \in \{-1,1\}^K} 2b^T y - b^T H b = 0. \tag{A.1}$$

Proposition 10 can be generalized [2], [18] to deal with arbitrary finite alphabets which are not part of the instance (and hence are fixed) of MULTIUSER DETECTION, i.e., the problem is inherently difficult when the number of users is large, regardless of the alphabet size. It is an open problem whether MULTIUSER DETECTION remains NP-hard when H is restricted to be Toeplitz. If this is the case, then it can be shown [18] that the problem of single-user maximum likelihood detection for intersymbol interference channels [17] is NP-hard in the length of the interference.

The usefulness and relevance of Proposition 10 stem from the fact that when the users are asynchronous, the cross correlations between their signals are unknown *a priori* and the worst-case computational complexity over all possible mutual offsets is the complexity measure of interest since it determines the maximum achievable data rate in the absence of synchronism among the users. Actually, no family of signature signals is known to result in optimum demodulation with polynomial-in-K complexity for all possible signal offsets. Thus even if the designer of the signal constellation were to include as a design criterion the complexity of the optimum demodulator in addition to the bit-error-rate performance (which dictates signals with low crosscorrelations), he would not be able to endow the signal set with any structure that would overcome the inherent intractability of the optimum asynchronous demodulation problem for all possible offsets.

The performance analysis of the optimum receiver for arbitrary energies and crosscorrelations is also inherently hard. According to (2.11) the maximum achievable asymptotic efficiency is obtained as the solution to multiuser asymptotic efficiency.

MULTIUSER ASYMPTOTIC EFFICIENCY—

Instance: Given $K \in Z^+$, $k \in \{1,\cdots,K\}$, and a nonnegative definite matrix $H \in Q^{K \times K}$;

Find: the kth user maximum asymptotic efficiency,

$$\eta_k = \frac{1}{w_k} \min_{\substack{\epsilon \in \{-1,0,1\}^K \\ \epsilon_k \neq 0}} \epsilon^T H \epsilon.$$

Proposition 11: MULTIUSER ASYMPTOTIC EFFICIENCY is NP-hard.

Proof: The proof is divided in two steps. First, $-1/0/1$ KNAPSACK is polynomially transformed to MULTIUSER ASYMPTOTIC EFFICIENCY. Then, $-1/0/1$ KNAPSACK is shown to be NP-complete. In analogy to the 0/1 KNAPSACK problem (e.g., [16]) we define

$-1/0/1$ KNAPSACK—

Instance: Given $L \in Z^+$, $G \in Z^+$ and a family of not necessarily distinct positive integers

$$\{l_i \in Z^+, i = 1,\cdots,L\};$$

Question: Are there integers $\epsilon_i \in \{-1,0,1\}$, $i = 1,\cdots,L$ such that $\Sigma_{i=1}^L \epsilon_i l_i = G$?

We transform $-1/0/1$ KNAPSACK to MULTIUSER ASYMPTOTIC EFFICIENCY by adding a user. Given $\{G, l_1,\cdots,l_L\}$, denote $l_{L+1} = G$ and construct the following instance: $K = L+1$, $k = L+1$, $h_{ij} = l_i l_j$, $1 \le i, j \le K$.

The Kth user asymptotic efficiency is equal to zero if and only if $\{G, l_1,\cdots,l_L\}$ is a "yes" instance of $-1/0/1$ KNAPSACK. To see this, note that we can fix $\epsilon_k = -1$ in the right side of (2.11)

[4] *Note added in proof:* This has now been accomplished in the companion paper [19], using a different approach.

without loss of generality. Then,

$$\eta_K = \frac{1}{G^2} \min_{\substack{\epsilon_i \in \{-1,0,1\} \\ 1 \le i \le K-1}} \left\{ h_{KK} + \sum_{n=1}^{K-1} \epsilon_n \left[-2h_{nK} + \sum_{m=1}^{K-1} \epsilon_m h_{nm} \right] \right\}$$

$$= \frac{1}{G^2} \min_{\substack{\epsilon_i \in \{-1,0,1\} \\ 1 \le i \le K-1}} \left(G - \sum_{n=1}^{K-1} \epsilon_n l_n \right)^2. \quad \text{(A.2)}$$

The proof that $-1/0/1$ KNAPSACK is NP-complete can be found in [18].

APPENDIX II

We give here an explicit procedure for finding the maximizing vector $\tilde{v}$ given implicitly by Proposition 6. The idea is the following: condition (4.19) states that if the maximizing vector $\tilde{v}$ lies in the intersection of a subset of the delimiting hyperplanes with equations $h_j^T\tilde{v} = 0$, $j \in S$, with S the index set of the specific hyperplanes, only the λ_j, $j \in S$ are possibly nonzero and enter into the expression defining $\tilde{v}$. Thus we have $|S|$ equations with $|S|$ unknowns, which we can solve to get the λ_i and then $\tilde{v}$. To state (and prove the correctness of) an algorithm that finds the optimum linear transformation, the following terminology is used.

Definition 1: Let S be an index set $\{j_1, j_2, \cdots, j_n\}$, $0 \le n \le K-1$, with $j_1, \cdots, j_n \in \{1, \cdots, K\} - \{k\}$, labeled in increasing order. Define

$$D_S(j) = \det \begin{vmatrix} h_j^T v_o & H_{jj_1} & \cdots & H_{jj_n} \\ h_{j_1}^T v_o & H_{j_1 j_1} & \cdots & H_{j_1 j_n} \\ \cdot & \cdot & \cdot & \cdot \\ h_{j_n}^T v_o & H_{j_n j_1} & \cdots & H_{j_n j_n} \end{vmatrix}. \quad \text{(A.3)}$$

Definition 2: We introduce an indicator for the second Kuhn–Tucker condition:

$$\text{if } e_j D_S(j) > 0, \text{ then } C_S(j) = \text{yes}, \qquad \text{else } C_S(j) = \text{no}. \quad \text{(A.4)}$$

Definition 3: An n-tuple S of $\{1, \cdots, K\} - \{k\}$ is *matched* if for all $i \in S$: $C_{S-\{i\}}(i) = \text{no}$.

Definition 4: An n-tuple S contains a basis B if $\{h_j | j \in B\}$ is a basis for $\{h_j | j \in S\}$.

Proposition 12: The following algorithm finds a vector $\tilde{v}$ satisfying (4.17)–(4.20).

A. Search for the index set with least cardinality $S \subseteq \{1, \cdots, K\} - \{k\}$, for which λ_i, $i \in S$, are possibly nonzero

```
n := 0
all n-tuples := untried; S_o := matched
WHILE n ≤ K − 2
    WHILE there is still an untried n-tuple containing a matched
    basis B
        select untried matched n-tuple := S_n, contained matched
        basis := B
        IF for all j ∉ S_n, j ≠ k, C_B(j) = yes, RETURN S_n, B, STOP
        ELSE S_n := tried
    RETURN
n := n + 1
RETURN
"decorrelating detector is optimal," output {2,···,K} − {k},
STOP.
```

B. Computation of the λ_i:

$i \notin B$: $\lambda_i = 0$

$i \in B$: λ_i are the solutions of the $|B|$ equations $|B|$ unknowns $h_i^T v = 0$, $i \in B$, where

$$v = v_o + \sum_{i \in B} \lambda_i e_i u_i.$$

C.

$$\tilde{v} = \frac{v_o + \sum_{i \in B} \lambda_i e_i u_i}{\left(v_o^T H v_o + v_o^T H \sum_{i \in B} \lambda_i e_i u_i \right)^{1/2}}.$$

Comment: Recall that this procedure has to be repeated for all the different $\{e_j\}$ in search of the maximal $\eta(e)$ value, until either the efficiency $\eta(e)$ reaches the upper bound given by the optimal detector, or all 2^K possibilities have been exhausted. Prior to running the algorithm, the sufficient conditions given in Propositions 8 and 9 should be checked.

Proof: Conditions (4.17) and (4.19) are obviously satisfied by construction of $\tilde{v}$ in C, and the requirement $h_j^T\tilde{v} = 0$ for the possibly nonzero λ_i in B. To prove conditions (4.18) and (4.20), consider the system of $|B|$ linear equations in $|B|$ unknowns of B. From A the set B is matched, and satisfies $C_B(j) = \text{yes}$ for all $j \neq k$, $j \notin S_n$. We have to show a) $\lambda_i \ge 0$, for all $i = 1, 2, \cdots, K$; and b) $C_B(j) = \text{yes}$ for all $j \neq k$, $j \notin S_n$ is equivalent to condition (4.18).

a) $\lambda_i = 0$, $i \notin B$, by construction of the index set S_n and B. For $i \in B$, in step B we solve $h_{j_i}^T\tilde{v} = 0$, all $i = 1, 2, \cdots, |B|$. Let $|B| = n$. Then

$$h_j^T v_o + \lambda_{j_1} e_{j_1} H_{j_1 j_1} + \cdots + \lambda_{j_n} e_{j_n} H_{j_1 j_n} = 0$$

$$h_{j_2}^T v_o + \lambda_{j_1} e_{j_1} H_{j_2 j_1} + \cdots + \lambda_{j_n} e_{j_n} H_{j_2 j_n} = 0$$

$$h_{j_n}^T v_o + \lambda_{j_1} e_{j_1} H_{j_n j_1} + \cdots + \lambda_{j_n} e_{j_n} H_{j_n j_n} = 0. \quad \text{(A.5)}$$

Denote by D_B the determinant of the coefficient matrix of the $\lambda_{j_i} e_{j_i}$. Since B is a basis and the corresponding matrix is nonnegative definite, D_B is strictly positive. Then, by Cramer's rule,

$$\lambda_{j_i} = \frac{-e_{j_i} D_{B-\{j_i\}}(j_i)}{D_B}. \quad \text{(A.6)}$$

The numerator is obtained by i row flips and i column flips to get j_i into position (1,1). Since the set B is matched, the numerator is nonnegative. As obtained above, the denominator is positive, hence $\lambda_i \ge 0$ for all $i \in B$. This completes the proof of a).

b) $h_j^T v = 0$, $j \in S_n$. For $j \notin S_n$, $j \neq k$, with the obtained values for λ compute the feasibility expressions:

$$e_j h_j^T \tilde{v} = e_j h_j^T \left(v_o + \sum_{i \in B} \lambda_i e_i u_i \right)$$

$$= \frac{e_j}{D_B} \left(D_B h_j^T v_o + \sum_{i \in B} - D_{B-\{i\}}(i) H_{ji} \right)$$

$$= \frac{1}{D_B} e_j D_B(j) > 0, \quad \text{(A.7)}$$

since $C_B(j) = \text{yes}$. The last equality is obtained by expanding along the first row of $D_B(j)$. This completes the proof of b). By construction the algorithm terminates after at most $K-2$ steps.

In part A of the algorithm notice that $n = 0$ corresponds to a solution in the interior of the feasible cone, with all λ equal to zero, and $\tilde{v} = v_o / \sqrt{v_o^T H v_o}$. The corresponding asymptotic efficiency $\eta^2(e)/w_k = v_o^T H v_o / w_k = \eta$, which is equal to the asymptotic efficiency of the maximum likelihood detector as given by (2.11). On the other hand, $n=1$ corresponds to a solution on exactly one of the delimiting hyperplanes, with exactly one λ nonzero (call it λ_j), and

$$\tilde{v}=\frac{1}{\eta(e)}\left(v_o-\frac{h_j^Tv_o}{H_{jj}}u_j\right) \tag{A.8}$$

and

$$\eta^2(e)=v_o^THv_o-\frac{\left(h_j^Tv_o\right)^2}{H_{jj}}. \tag{A.9}$$

The asymptotic efficiency achieved in this case is bounded above by the one for $n=0$, since the second term is nonnegative. If the matrix H does not have a lot of structure, which is to be expected in practical applications, this is the most probable case. For increasing n the computational effort grows fast, but in most cases the algorithm will terminate for very small n.

We also have an explicit solution for the "terminal case," $n=K-1$, which corresponds to the decorrelating detector case. Then, without loss of generality, $\tilde{v}=h_k^I/\sqrt{H_{kk}^I}$, a scaled version of the kth column of any generalized inverse matrix of H (in particular of H^+) and $\eta(e)=1/H_{kk}^+$, which is equal to the kth user asymptotic efficiency of the decorrelating detector, when the scaling factor $1/w_k$ of (4.16) is taken into account. This can be shown as follows. In the terminal case $h_j^T\tilde{v}=0$, for all $j\neq k$. Hence

$$\eta(e)=\max_{\substack{v\in R^K\\ cHv=1\\ h_j^Tv=0\\ j\neq k}} v_o^THv=\max_{\substack{v\in R^K\\ h_j^Tv=0\\ j\neq k\\ v_kh_k^Tv=1}} h_k^Tv=\max_{\substack{v\in R^K\\ Hv=(1/v_k)u_k}} \frac{1}{v_k}. \tag{A.10}$$

If user k is dependent, $Hv=0$ and $\eta(e)=0$. Since this was the best choice of v, we can without loss of generality replace $\tilde{v}$ by the kth row of any generalized inverse, because the resulting asymptotic efficiency cannot become negative. If user k is independent, Lemma 1 implies $HH^Iu_k=u_k$, and for all v in the feasible set,

$$H^IHv=\frac{1}{v_k}H^Iu_k.$$

Hence using Lemma 1 for both equations, we obtain

$$v_k=\sqrt{H_{kk}^I}=\sqrt{H_{kk}^+}.$$

The feasible set in (A.10), $F=\{v|Hv=(1/v_k)u_k\}$, is nonempty (e.g., it contains the set $\{(1/v_k)h_k^I, H^I\in I(H)\}$), and for all $v\in F$, $v_k=\sqrt{H_{kk}^+}$. Hence $\eta(e)=1/\sqrt{H_{kk}^+}$, and with (3.8), $(1/w_k)\eta^2(e)=1/R_{kk}^+$, which is the energy independent asymptotic efficiency of the decorrelating detector for independent users.

References

[1] S. Verdú, "Minimum probability of error for asynchronous Gaussian multiple-access channels," *IEEE Trans. Inform. Theory*, vol. IT-32, pp. 85–96, Jan. 1986.

[2] ——, "Optimum multi-user signal detection," Ph.D. dissertation, Dept. Elec. Comput. Eng., Univ. of Illinois, Urbana-Champaign, Coordinated Sci. Lab., Urbana, IL, Rep. T-151, Aug. 1984.

[3] ——, "Optimum multi-user asymptotic efficiency," *IEEE Trans Commun.*, vol. COM-34, no. 9, pp. 890–897, Sept. 1986

[4] I. N. Bronshtein and K. A. Semendyayev, *Handbook of Mathematics*. Verlag Harri Deutsch, 1985.

[5] J. Wozencraft and I. Jacobs, *Principles of Communication Engineering*. New York: Wiley, 1965.

[6] K. S. Schneider, "Optimum detection of code division multiplexed signals," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-15, pp. 181–185, Jan. 1979.

[7] T. Kailath, *Linear Systems*. Englewood Cliffs, NJ. Prentice-Hall, 1980.

[8] R. Lupas-Golaszewski and S. Verdú, "Asymptotic efficiency of linear multiuser detectors," in *Proc. 25th IEEE Conf Decision and Control*, Athens, Greece, Dec. 1986, pp. 2094–2100.

[9] D. Luenberger, *Optimization by Vector Space Methods*. New York. Wiley, 1969.

[10] T. L. Boullion and P. L. Odell, *Generalized Inverse Matrices*. New York: Wiley-Interscience, 1971.

[11] P. Lancaster and M. Tismenetsky, *The Theory of Matrices*. New York: Academic, 1985.

[12] H. V. Poor, *An Introduction to Signal Detection and Estimation*. New York: Springer-Verlag, 1988.

[13] C. W. Helstrom, *Statistical Theory of Signal Detection*, 2nd ed. London: Pergamon, 1968.

[14] S. Verdú, "Optimum multi-user detection is NP-hard," presented at the 1985 Int. Symp. Information Theory, Brighton, UK, June 1985.

[15] M. R. Garey and D. S. Johnson, *Computers and Intractability. A Guide to the Theory of NP-Completeness*. San Francisco, CA. Freeman, 1979.

[16] C. H. Papadimitriou and K. Steiglitz, *Combinatorial Optimization Algorithms and Complexity*. Englewood Cliffs, NJ: Prentice-Hall, 1982.

[17] G. D. Forney, "Maximum likelihood sequence estimation of digital sequences in the presence of intersymbol interference," *IEEE Trans. Inform. Theory*, vol. IT-18, no. 3, pp. 363–378, May 1972.

[18] S. Verdú, "Computational complexity of optimum multiuser detection," *Algorithmica*, vol. 4, 1989, to appear.

[19] R. Lupas and S. Verdú, "Near-far resistance of multiuser detectors in asynchronous channels," *IEEE Trans. Commun.*, vol. 37, 1989, to appear.

Algorithmica (1989) 4: 303-312

Algorithmica

Computational Complexity of Optimum Multiuser Detection[1]

Sergio Verdú[2]

Abstract. Optimum centralized demodulation of the independent data streams transmitted simultaneously by several users through a Code Division Multiple-Access channel is considered. Each user sends an arbitrary assigned signal waveform, which is linearly modulated by symbols drawn from a finite alphabet. If the users are asynchronous, the optimum multiuser detector can be implemented by a Viterbi algorithm whose time-complexity is linear in the number of symbols transmitted by each user and exponential in the number of users. It is shown that the combinatorial problem of selecting the most likely transmitted data stream given the sufficient statistics (sequence of matched filter outputs), and the signal energies and cross-correlations is nondeterministic polynomial-time hard (NP-hard) in the number of users. And it remains so even if the users are restricted to be symbol-synchronous.

The performance analysis of optimum multiuser detection in terms of the set of multiuser asymptotic efficiencies is equivalent to the computation of the minimum Euclidean distance between any pair of distinct multiuser signals. This problem is also shown to be NP-hard and a conjecture on a longstanding open problem in single user data communication theory is presented.

Key Words. NP-complete, Hypothesis testing, Code Division Multiple Access, Gaussian communication channels, Maximum-likelihood sequence detection.

1. Introduction. The purpose of hypothesis testing problems is to select a solution (decision) from among a *finite* set of possible solutions (hypotheses). Typically, the number of hypotheses is small, in which case the inherent combinatorial optimization nature of the problem does not play any role and the main question is to obtain the values of the likelihood function or other finite-dimensional set of sufficient statistics. In this paper we study a data demodulation problem where the reverse situation is encountered: it is straightforward to obtain a set of scalar sufficient statistics but the number of hypotheses is very large.

An important problem arising in multipoint-to-point digital communication networks (e.g., radio networks, local-area networks, and uplink satellite channels) is the optimum centralized demodulation of the information sent simultaneously by several users through a Gaussian multiple-access channel. Even though the users may not employ a protocol to coordinate their transmission epochs, effective sharing of the channel is possible because each user modulates a different signature signal waveform which is known by the intended receiver (Code Division Multiple Access (CDMA)). Recently [1], optimum multiuser detection has been shown to offer important gains in bit-error-rate performance over single-user detectors, which are conventionally used in practice and neglect the

[1] This work was partially supported by the National Science Foundation under Grant ECS-8504752, and by the US Army Research Office under Contract DAAL03-87-K-0062.
[2] Department of Electrical Engineering, Princeton University, Princeton, NJ 08544, USA.

Received December 17, 1986; revised June 1, 1988. Communicated by Israel Cidon and Inder S. Gopal.

presence of interfering users. The optimum multiuser receiver can be viewed as a bank of single-user detectors followed by a common algorithm that selects the most likely transmitted symbols. The structure of this algorithm depends crucially on whether or not the users maintain symbol synchronism. In the synchronous case, it is enough to maximize a quadratic function, while in the asynchronous case the way to get rid of the interference among the users is to employ a version of the Viterbi forward dynamic programming algorithm [2], whose time-complexity is exponential in the number of active users. It is shown in this paper that the problem is nondeterministic polynomial-time hard (NP-hard) in the number of users, and hence there exists no polynomial-time algorithm for optimum multiuser detection unless such an algorithm is found for a large class of combinatorial problems, such as the traveling salesman and integer linear programming problems. It is shown that the problem remains NP-hard even in the synchronous case despite an earlier claim of existence of polynomial solutions [3] for that case.

In Section 2 the multiple-access channel model and the maximum-likelihood detection problem are formulated and it is shown that optimum multiuser detection is NP-hard in the number of users. In Section 3 it is shown that the performance analysis of optimum multiuser detectors is intrinsically difficult due to the fact that the computation of the minimum distance between any pair of distinct multiuser signals is also NP-hard. Finally, Section 4 summarizes the main points of the paper, discusses suboptimum alternatives, and presents a conjecture on a longstanding open problem in data communication theory.

2. Optimum Multiuser Detection. Assume each of K users transmits independent symbols by modulating a preassigned waveform from a signal constellation $\{s_k(t),\ t\in[0, T],\ k=1,\dots,K\}$. If the users cooperate to maintain symbol synchronism, the receiver observes the sum of the modulated signals imbedded in noise, i.e.,

$$r(t)=\sum_{k=1}^{K} b_k s_k(t)+n(t), \tag{1}$$

where the symbols b_k, $k=1,\dots,K$, are drawn by each user from a finite alphabet A. A reasonable decision rule is to select the set of symbols corresponding to that signal among the possible ones that resembles most closely (in a mean-square sense) the received waveform. If the noise is Gaussian and white, then this rule is optimum in the maximum-likelihood sense. If, furthermore, all vectors $\mathbf{b}=(b_1,\dots,b_K)\in A^k$ are *a priori* equiprobable, then the minimum distance rule gives the maximum-*a-posteriori* (MAP) decision. In the single-user case, this detector is implemented by comparing the output of a matched filter with a set of thresholds. Analogously, in the multiuser problem we have

$$\arg\min_{b\in A^K}\left\|r(t)-\sum_{k=1}^{K} b_k s_k(t)\right\|=\arg\max_{b\in A^K} 2\mathbf{b}^T\mathbf{y}-\mathbf{b}^T\mathbf{H}\mathbf{b}, \tag{2}$$

where $\mathbf{y}=(y_1,\ldots,y_K)$, $y_k=\int_0^T s_k(t)r(t)\,dt$, i.e., y_k is the output of the matched filter of the kth user signal and the entries of the nonnegative definite matrix $\mathbf{H}$ are given by

$$h_{ij}=\int_0^T s_i(t)s_j(t)\,dt. \tag{3}$$

If the signal waveforms are orthogonal, then $\mathbf{H}$ is a diagonal matrix and the maximization in (2) decouples into K single-user problems. Otherwise (in practice, there may be bandwidth or complexity constraints that prevent the designer from choosing an orthogonal signal set), a combinatorial algorithm is required to solve the quadratic optimization (2) over the finite set A^K, given the vector of sufficient statistics $\mathbf{y}$ and the signal cross-correlations $\mathbf{H}$. Since the set of quantities $\{\mathbf{b}^T\mathbf{y}, \mathbf{b}\in A^K\}$ can be computed in $O(|A|^K)$ operations, an upper bound on the time-complexity per bit (TCB)[3] required to solve (2) is $O(|A|^K/K\log|A|)$. This is the best available upper bound; in [3] it is claimed that a receiver whose complexity is polynomial in the number of users (basically, if $A=\{-1,+1\}$, select the sign of the components of $\mathbf{H}^{-1}\mathbf{y}$) is optimal. Unfortunately, this claim is erroneous; the mistake in the derivation of the detector is commited in equation 4 of [3] where it is implicitly assumed that the symbols put out by the detector are uncorrelated with the noise component of the matched filter outputs.

More significant in practical applications is the case where the users are mutually asynchronous, and indeed one of the chief advantages of CDMA over other channel sharing strategies is that no type of coordination among the users is required. Now, however, (1) is no longer a valid model. The delays $\{\tau_k, k=1,\ldots,K\}$ account for the offsets between the signaling epochs and (1) has to be generalized to

$$r(t)=\sum_{i=-M}^{M}\sum_{k=1}^{K} b_k(i)s_k(t-iT-\tau_k)+n(t), \tag{4}$$

where by convention $s_k(t)=0$ for $t\notin[0,T]$. Now we can no longer restrict our attention to the one-shot case because optimum decisions are based on the whole received waveform due to the interference between the symbols. The optimum receiver [1] for the asynchronous case in the sense of selecting the most likely sequence of symbols consists of a front-end of matched filters (just as in the synchronous case) followed by a Viterbi dynamic programming algorithm with $|A|^{K-1}$ states and a periodically time-varying branch metric. The TCB of this decision algorithm is $O(|A|^K/\log|A|)$, and hence the penalty in time-complexity due to the lack of synchronism between the users is slight. The usefulness and relevance of the computational complexity results proved in this paper stem from the fact that when the users are asynchronous, the cross-correlations between

[3] The time-complexity per bit is defined as the limit of the ratio of total time to the number of demodulated bits as this goes to infinity. Note that any preprocessing of the signal cross-correlations does not affect TCB.

their signals are unknown *a priori*, and the worst-case TCB over all possible mutual offsets is the complexity measure of interest since it determines the maximum achievable data rate in the absence of synchronism among the users. Actually, no family of signature signals is known to result in optimum demodulation with polynomial-in-K complexity for all possible signal offsets. So even if the designer of the signal constellation were to include in his design criterion the complexity of the optimum demodulator in addition to the bit-error-rate performance (which dictates signals with low cross-correlations), he would not be able to endow the signal set with any structure that would overcome the inherent intractability of the optimum asynchronous demodulation problem for all possible offsets. On the other hand, in the synchronous case, the designer of the signal constellation has more control on the cross-correlations (subject to constraints such as bandwidth or number of chips per symbol in Direct-Sequence Spread-Spectrum), and it is conceivable that there exist synchronous signal design constraints that result in families of signature signals whose structure can be exploited to result in optimum polynomial-time decision algorithms. This is the reason why the computational complexity results of this paper appear to be more relevant to asynchronous channels even though for the purposes of the proofs in our lower bound analysis we may restrict attention to the special case where all the delays coincide ($\tau_1 = \cdots = \tau_K$), because the optimum multiuser detector must be able to deal with any arbitrary set of delays.

In order to ascertain that the intractability of the optimum multiuser problem arises when the number of users is large and the alphabet size is kept constant, we first fix an arbitrary alphabet $A = \{a_1, \ldots, a_m\}$ (which is a set of integers satisfying $a_i < a_{i+1}$), and define a class of instances of the combinatorial optimization problem for that *fixed* A.[4]

MULTIUSER DETECTION
Instance: Given $K \in \mathbb{Z}^+$, $\mathbf{y} \in Q^K$, and a nonnegative definite matrix $\mathbf{H} \in Q^{K\times K}$
Find: $\{\mathbf{b}^* \in A^K\}$ that maximizes $2\mathbf{b}^T\mathbf{y} - \mathbf{b}^T\mathbf{H}\mathbf{b}$.

PROPOSITION 1. *If* $|A| > 1$, *then MULTIUSER DETECTION is* NP-*hard.*

PROOF. The proof of NP-hardness of MULTIUSER DETECTION can be carried out by transformation from PARTITION, an NP-complete recognition problem. Recall its definition [4]:

Instance: Given $L \in \mathbb{Z}^+$ and $\{l_i \in \mathbb{Z}^+, i = 1, \ldots, L\}$.
Question: Is there a subset $I \subset \{1, \ldots, L\}$ such that $\sum_{i \in I} l_i = \sum_{i \notin I} l_i$?

For each instance of PARTITION, we can find in polynomial time an instance of MULTIUSER DETECTION whose solution can in turn be processed in

[4] Note that since the alphabet A is not part of the instance, if a specific A is assumed, then the corresponding NP-hard result is a *corollary* to Proposition 1. Actually, for $A = \{-1, +1\}$, the proof of Proposition 1 can be simplified considerably by letting $h_{ij} = l_i l_j$ and $y_k = 0$ therein.

polynomial time to give an answer to PARTITION. Given $l_1, \ldots, l_L$, we choose the following instance of MULTIUSER DETECTION:

$$K = L,$$

$$h_{ij} = l_i l_j, \qquad i \neq j,$$

$$h_{ii} = \left\lceil l_i \max\left\{ l_i, (a_2 - a_1)^{-1}[2a_m - a_1 - a_2] \sum_{\substack{j=1 \\ j \neq i}}^{K} l_j \right\} \right\rceil, \qquad i = 1, \ldots, K,$$

$$y_k = \tfrac{1}{2}(a_1 + a_2)\left(h_{kk} + l_k \sum_{\substack{j=1 \\ j \neq k}}^{K} l_j \right), \qquad k = 1, \ldots, K.$$

Note that this is a valid instance of MULTIUSER DETECTION, because $\mathbf{H}$ is a nonnegative definite matrix. Once the solution to this instance of MULTIUSER DETECTION is found, we can find the solution to the original instance of PARTITION, because $\{l_1, \ldots, l_L\}$ is a *yes* instance of PARTITION if and only if

$$\max_{b \in A^K} 2\mathbf{b}^T\mathbf{y} - \mathbf{b}^T\mathbf{H}\mathbf{b} = \left[\tfrac{1}{2}(a_1 + a_2) \sum_{i=1}^{K} l_i \right]^2 + a_1 a_2 \sum_{i=1}^{K} (h_{ii} - l_i^2). \tag{5}$$

Equation (5) can be shown by changing the variable in the left-hand side of (5) $\mathbf{b} = \frac{1}{2}(a_2 - a_1)\mathbf{z} + \frac{1}{2}(a_1 + a_2)$ and proving that it is enough to restrict attention to the values $z_i = \pm 1$ in the maximization in (5) (see [5] for details). □

The foregoing proof shows that the same transformation works if the value of the diagonal elements of $\mathbf{H}$ is arbitrarily increased. Hence MULTIUSER DETECTION remains NP-hard if $\mathbf{H}$ is restricted to be strongly diagonal (an important special case in CDMA with equal-energy users). Note that MULTIUSER DETECTION was defined for a fixed arbitrary alphabet. Thus, Proposition 1 implies that the problem is inherently difficult when the number of users is large, regardless of the alphabet size (often a small integer). Conversely, in Section 2 we saw that the problem is polynomial in the alphabet size for fixed number of

3. NP-hardness of Multiuser Asymptotic Efficiency. In this section we examine the complexity of the *performance analysis* of optimum multiuser detection. The purpose of this analysis is to evaluate the effect of the energies and cross-correlations of the signal constellation on the bit-error-rate of the receiver for an arbitrary level of background noise. It has been shown [6] that the key performance measure is the multiuser asymptotic efficiency, or ratio between the exponential decay rate of the bit-error-rates with and without interfering users. This parameter effectively quantifies the degradation in bit-error-rate due to the presence of other users, in situations where the background noise is not dominant. The asymptotic efficiency of the kth user, η_k, is proportional to the Euclidean distance between any pair of transmitted signals whose kth symbols do not agree [6]. Specifically, assuming synchronous users and antipodal modulation (i.e., $A=\{-1,+1\}$), the kth user asymptotic efficiency can be expressed as

$$\eta_k = \frac{\min_{b_k^1 \neq b_k^2} \int_0^T \left[\sum_{i=1}^{K} (b_i^1 - b_i^2) s_i(t)\right]^2 dt}{4\int_0^T s_k^2(t)\,dt} = \frac{1}{h_{kk}} \min_{\substack{\varepsilon \in \{-1,0,1\}^K \\ \varepsilon_k \neq 0}} \varepsilon^T \mathbf{H} \varepsilon. \tag{6}$$

PROPOSITION 2. *The following problem is* NP-*hard*:

MULTIUSER ASYMPTOTIC EFFICIENCY
Instance: *Given* $K \in \mathbb{Z}^+$, $k \in \{1, \dots, K\}$, *and a nonnegative matrix* $\mathbf{H} \in \mathbb{Z}^{K \times K}$
Find: *The* kth *user maximum asymptotic efficiency,*

$$\eta_k = (1/h_{kk}) \min_{\varepsilon \in \{-1,0,1\}^K, \varepsilon_k \neq 0} \varepsilon^T \mathbf{H} \varepsilon.$$

PROOF. The proof is divided in two steps, first $-1/0/1$ KNAPSACK is polynomially transformed to MULTIUSER ASYMPTOTIC EFFICIENCY, and then we show that $-1/0/1$ KNAPSACK is NP-complete. In analogy to the 0/1 KNAPSACK problem (e.g., [7]) we define

$-1/0/1$ KNAPSACK
Instance: $L \in \mathbb{Z}^+$, $G \in \mathbb{Z}^+$, and a family of not necessarily distinct positive integers $\{l_i \in \mathbb{Z}^+, i = 1, \dots, L\}$.
Question: Are there integers $\varepsilon_i \in \{-1, 0, 1\}$, $i = 1, \dots, L$, such that $\sum_{i=1}^{L} \varepsilon_i l_i = G$?

We transform $-1/0/1$ KNAPSACK to MULTIUSER ASYMPTOTIC EFFICIENCY by adding an additional user. Given $\{G, l_1, \dots, l_L\}$, denote $l_{L+1} = G$ and construct the following instance: $K = L+1$, $k = L+1$, $h_{ij} = l_i l_j$, $1 \leq i, j \leq K$.

The Kth-user asymptotic efficiency is equal to 0 if and only if $\{G, l_1, \dots, l_L\}$ is a *yes* instance of $-1/0/1$ KNAPSACK. To see this, note that we can fix $\varepsilon_k = -1$ in the right-hand side of (6) without loss of generality. Then,

$$\begin{aligned} \eta_K &= \frac{1}{G^2} \min_{\substack{\varepsilon_i \in \{-1,0,1\} \\ 1 \leq i \leq K-1}} \left\{ h_{KK} + \sum_{n=1}^{K-1} \varepsilon_n \left[-2h_{nK} + \sum_{m=1}^{K-1} \varepsilon_m h_{nm} \right] \right\} \\ &= \frac{1}{G^2} \min_{\substack{\varepsilon_i \in \{-1,0,1\} \\ 1 \leq i \leq K-1}} \left(G - \sum_{n=1}^{K-1} \varepsilon_n l_n \right)^2. \end{aligned} \tag{7}$$

Now we show that $-1/0/1$ KNAPSACK is NP-complete. Its membership in NP is obvious. Note that it is easy to transform $-1/0/1$ KNAPSACK to 0/1 KNAPSACK ($\{G, l_1, \dots, l_L\}$ is a *yes* instance of $-1/0/1$ KNAPSACK if and only if $\{G, l_1, -l_1, \dots, l_L, -l_L\}$ is a *yes* instance of 0/1 KNAPSACK). However, we need to show the reverse transformation, namely, fixing any instance of 0/1

KNAPSACK obtain an equivalent instance of $-1/0/1$ KNAPSACK. The idea is reminiscent of the polynomial transformation of 0/1 KNAPSACK to POSITIVE INTEGER KNAPSACK (see p. 376 of [7]), and it consists of constructing an augmented instance of $-1/0/1$ KNAPSACK whose integers are large enough to force every coefficient to be either 0 or 1. Choose an instance $\{G, l_1, \ldots, l_L\}$ of 0/1 KNAPSACK, and construct the following instance of $-1/0/1$ KNAPSACK $\{D, p_1, \ldots, p_{2L}\}$:

$$D = G + \sum_{j=1}^{L} M^j,$$

$$p_j = \begin{cases} l_j + M^j & \text{for } 1 \le j \le L, \\ M^{j-L} & \text{for } L+1 \le j \le 2L, \end{cases}$$

$$M = 1 + \max\left\{3, G + \sum_{i=1}^{L} l_i\right\}.$$

Then it follows from this choice that for all $\{\varepsilon_i \in \{-1, 0, 1\}, 1 \le i \le 2L\}$

$$\sum_{i=1}^{2L} \varepsilon_i p_i - D = -G + \sum_{i=1}^{L} \varepsilon_i l_i + \sum_{i=1}^{L} M^i(\varepsilon_i + \varepsilon_{i+n} - 1). \tag{8}$$

It is straightforward to show that M is too large to be a root of any L-degree polynomial $\sum_{i=0}^{L} \beta_i x^i$, where $\beta_i \in \{-3, -2, -1, 0, 1\}$ for $1 \le i \le L$ and $\beta_0 \in \{-G + \sum_{i=1}^{L} \varepsilon_i l_i, \varepsilon_i \in \{-1, 0, 1\}, 1 \le i \le L\}$. Hence, (8) is equal to zero if and only if all the coefficients of the polynomial in M of the right-hand side are zero, i.e.,

$$\sum_{i=1}^{2L} \varepsilon_i p_i = D \iff \sum_{i=1}^{L} \varepsilon_i l_i = G \quad \text{and} \quad \{\varepsilon_i = 0, \varepsilon_{i+n} = 1 \text{ or } \varepsilon_i = 1, \varepsilon_{i+n} = 0, 1 \le i \le l\}.$$

Therefore, $\{G, l_1, \ldots, l_L\}$ is a *yes* instance of 0/1 KNAPSACK if and only if $\{D, p_1, \ldots, p_{2L}\}$ is a *yes* instance of $-1/0/1$ KNAPSACK. □

Note that the proof of Proposition 2 shows in fact that the problem of deciding whether a signal constellation is uniquely decodable, i.e., whether different transmitted bits result in the same waveform, is NP-complete.

4. Concluding Remarks, Suboptimum Algorithms, and Open Problems. It has been shown that the problem of optimum detection in Gaussian multiple-access channels is NP-hard in the number of users. This result holds for any nontrivial alphabet even if the channel is symbol-synchronous. Exponential-in-K optimum detectors for Poisson multiple-access channels with point-process observations were obtained in [8]. It can be shown that this problem is also NP-hard in both the additive-rate and additive-light models of the channel.

Not only is the optimum decision rule intrinsically difficult, but so is the analysis of its performance due to the NP-hardness of the computation of multiuser asymptotic efficiencies. It should be pointed out, however, that in many instances of signal constellations used in CDMA Direct Sequence Spread-Spectrum systems [9], the cross-correlations are low enough to pass sufficient conditions [6] ensuring unit asymptotic efficiency that are computable in quadratic time in K. So, unlike the situation we encountered in the optimum asynchronous demodulation problem, the worst-case complexity measure may be overly pessimistic for specific instances exhibiting low cross-correlations.

What alternatives, then, does the designer have when the number of users is large? The suboptimum solution currently employed in practice is the bank of single-user receivers (i.e., a matched filter for each user followed by a threshold). Unfortunately, this scheme achieves far from optimum bit-error-rate and its performance breaks down when the signal energies are dissimilar (the near-far problem) [1], [6], [10]. Therefore, the search for approximation algorithms that achieve near optimum bit-error-rate with polynomial complexity appears to be an open research area with important consequences in practice. Numerical results indicate that performance extremely close to the single-user lower bound is achievable in the bit-error-rate region of usual interest (10^{-4} or less) by the maximum-likelihood multiuser receiver even if the cross-correlation qualities of signal constellations typically used in practice are considerably relaxed. This is a sign that the Viterbi-based optimum multiuser receiver possesses an important degree of redundancy in situations with good signal sets and low background noise and hence faster decision algorithms achieving similar performance are plausible. Furthermore, a linear multiuser demodulator whose TCB is linear in K is found in [10] to achieve the same worst-case asymptotic efficiency over the energies of the interfering users[5] as the optimum demodulator. While for specific values of the received energies, its asymptotic efficiency (and hence, its bit error rate) need not be close to the optimum one, its performance is guaranteed to exceed a high lower bound in all cases of practical interest, thus making this suboptimum demodulator an attractive choice from both the complexity and performance standpoints.

Finally, let us consider an interesting special case of the asynchronous optimum multiuser detection model in (4), namely the single-user intersymbol interference problem,

$$r(t)=\sum_{i=-M}^{M} b(i)s(t-iT)+n(t),$$

where the duration of $s(t)$ is greater than T. In general, there is no known efficient method to obtain the most likely sequence of transmitted symbols given the received waveform (TCB is exponential in the frame length M). However, if the number of signals that interfere at any given time is bounded by, say, L, then it

[5] This is called the *near-far resistance*, a key measure of the robustness of the system against variations in the received energies.

is well known [11] that maximum-likelihood sequence detection can be implemented by the Viterbi algorithm in TCB which although exponential in L is independent of M. Despite many efforts (e.g., [12]-[14]) motivated by the importance of this problem in the area of data transmission through bandlimited channels, no polynomial-in-L algorithm for maximum-likelihood sequence detection is known. This fact and the results of this paper lead us to suspect that we may be facing another NP-hard problem. In fact following the same steps as in Section 2, it can be seen that the most likely sequence of symbols corresponding to (13) is the one that maximizes $2\mathbf{b}^T\mathbf{y}-\mathbf{b}^T\mathbf{H}\mathbf{b}$, where $\mathbf{b}=(b(-M),\ldots,b(M))$, $\mathbf{y}=(y(-M),\ldots,y(M))$, $y(i)=\int s(t-iT)r(t)\,dt$, and $\mathbf{H}$ is the nonnegative definite Toeplitz matrix (i.e., constant along diagonals) with entries given by $h_{ij}=\int s(t-iT)s(t-jT)\,dt$. Hence if we specialize $L=2M+1=K$, the underlying combinatorial problem coincides with MULTIUSER DETECTION with an additional restriction on the data:

CONJECTURE 1. MULTIUSER DETECTION remains NP-hard if $\mathbf{H}$ is restricted to be Toeplitz.

Indeed, it appears that the Toeplitz condition imposes an analytically inconvenient restriction on the set of allowable instances. A possible route is to consider the following restricted version of the problem.

FIR
Instance: Given $L\in\mathbb{Z}^+$, $E\in\mathbb{Z}^+$, and the coefficients of a finite-impulse response (FIR) digital filter of length L ($h_i\in\mathbb{Z}$, $i=0,\ldots,L-1$).
Question: Does there exist an input sequence ($b_i\in\{-1,+1\}$, $i=0,\ldots,L-1$) such that the output energy of the FIR is less than E, i.e.,

$$\sum_{i=0}^{2L-2}\left(\sum_{j=0}^{L-1} b_j h_{i-j}\right)^2 \le E?$$

CONJECTURE 2. FIR is NP-complete.

Similarly, the problem of the performance analysis of the single-user intersymbol interference channel is equivalent to finding the minimum distance between any pair of transmitted data streams. This problem for which no polynomial algorithm in the length of the interference is known can be put as a special case of MULTIUSER ASYMPTOTIC EFFICIENCY and it is not known whether it is NP-hard.

References

[1] S. Verdú, Minimum Probability of Error for Asynchronous Gaussian Multiple-Access Channels, *IEEE Trans. Inform. Theory*, 32, pp. 85-96 (January 1986).
[2] G. D. Forney, The Viterbi Algorithm, *Proc. IEEE*, 6?, pp. 268-278 (March 1973).
[3] K. S. Schneider, Optimum Detection of Code Division Multiplexed Signals, *IEEE Trans. Aerospace Electron. Systems*, 15, pp. 181-185 (January 1979).

[4] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-completeness*, Freeman, San Francisco (1979).

[5] S. Verdú, Optimum Multi-user Signal Detection, Ph.D. Dissertation, Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign. Report T-151, Coordinated Science Laboratory, Urbana, IL (August 1984).

[6] S. Verdú, Optimum Multiuser Asymptotic Efficiency, *IEEE Trans. Comm.*, **34**, pp. 890-897 (September 1986).

[7] C. H. Papadimitrou and K. Steiglitz, *Combinatorial Optimization: Algorithms and Complexity*, Prentice-Hall, Englewood Cliffs, NJ (1982).

[8] S. Verdú, Multiple-Access Channels with Point-Process Observations: Optimum Demodulation, *IEEE Trans. Inform. Theory*, **32**, pp. 642-651 (September 1986).

[9] M. K. Simon, J. K. Omura, R. A. Scholtz, and B. K. Levitt, *Spread Spectrum Communications*, Vol. 3, Computer Science Press, Rockville, MD (1985).

[10] R. Lupas and S. Verdú, Linear Multiuser Detectors for Synchronous Code-Division Multiple-Access Channels, *IEEE Trans. Inform. Theory*, **35** (January 1989).

[11] A. J. Viterbi and J. K. Omura, *Principles of Digital Communication and Coding*, McGraw-Hill, New York (1979).

[12] G. J. Foschini, A Reduced State Variant of Maximum Likelihood Sequence Detection Attaining Optimum Performance for High Signal-to-Noise Ratios, *IEEE Trans. Inform. Theory*, **23**, pp. 605-609 (September 1977).

[13] A. P. Clark, *Advanced Data-Transmission Systems*, Halsted Press, New York (1977).

[14] D. Kazakos, Computational Savings and Implementation of Maximum Likelihood Detectors, *IEEE Trans. Inform. Theory*, **24**, pp. 124-126 (January 1978).

A Semi-Classical Analysis of Optical Code Division Multiple Access

David Brady, Sergio Verdú †

Department of Electrical Engineering
Princeton University,
Princeton, NJ 08544

ABSTRACT

In this paper we consider a noncoherent, optical, asynchronous, code division multiple access (CDMA) system. We present a semi-classical analysis of the error rate for a single-user, matched-filter receiver that applies for arbitrary photomultipliers and signature sequence sets, adheres fully to the semi-classical model of light, and does not depend on approximations for large user groups, strong received optical fields, or chip synchronism. We compare the exact minimum probability of error and optimal threshold to those obtained with popular approximations on user synchronism or on the distribution of the multiple access interference (MAI). For the special case of unity-gain photodetectors and prime sequences, we show that the approximation of chip synchronism yields a weak upper bound on the exact error rate. We demonstrate that the approximations of perfect optical-to-electrical conversion and Gaussian-distributed MAI yield a poor approximation to the minimum error rate and an underestimate of the optimal threshold. In this paper we also develop arbitrarily-tight bounds on the error rate for unequal energies per bit. In the case when the signal energies coincide, these bounding expressions are considerably easier to compute than the exact error rate.

1. Introduction

Several users may independently access a common communication channel using code division multiple access (CDMA) modulation. This multiaccess scheme does not use time- or frequency-allocation, and users may transmit without the delays inherent to multiaccess protocols. In the direct-detection optical CDMA channel, interference immunity is achieved by the assignment of rapidly varying, on-off waveforms. These waveforms, or signature sequences, are modulated by the data of each user and concentrate the transmitted energy into relatively short time intervals in each symbol period. The transmitted signals from the users are then combined on a common optical fiber. Single-user demodulation is (suboptimally) achieved by correlating the aggregate signal and the signature sequence of the desired user. As a result, the correlator output is the desired signal in additive interference, which is reduced through the use of signature-sequence sets with low cross-correlation. A correlator receiver of this type must know only the timing epoch of the desired user, and a common timing reference need not be sent to all transmitters.

In this paper we present the error rate of a particular single-user receiver in the noncoherent CDMA optical fiber channel. This receiver has been the focus of previous analyses and local area network prototypes [1-3]. In contrast to previous efforts we have avoided making approximations to three analytical obstacles. First, our analysis retains the quantized nature of electromagnetic radiation. While the particle nature of radiation may be neglected for the analysis of microwave communication systems, optical

† This work was supported by the U.S. Army Research Office under Contract DAAL03-87-K-0062.

illumination of a surface is most accurately described by a photon arrival process, a doubly-stochastic point process whose random rate is proportional to the intensity of the electromagnetic field integrated over the illuminated surface. In particular, this holds for field intensities typical of received signals in direct-detection optical systems [4-6]. As might be expected, it is convenient to analyze receivers based on observations of the received-field intensity, rather than on observations of a point process driven by this intensity. This approximation is equivalent to neglecting the quantization of energy at the photodetector screen, and is increasingly inaccurate as the received optical power decreases. Second, we have avoided the use of a Gaussian approximation of the MAI distribution. Due to low-weight signature sequences, the MAI is additively composed of cross-correlations that are usually limited to a few chips of coincidence, and its maximum is too small to apply central-limit-theorem arguments. As might be expected, the derivation of the optimal hypothesis test and the error analysis are complicated by the nature of the MAI distribution and would be simplified if the MAI had Gaussian statistics. Third, we have avoided a chip-synchronous approximation on the relative delays between the users, in which case the relative delays are integer multiples of the chip interval. Since CDMA is a form of multiuser modulation without transmission coordination, the relative delay between any two users is uniformly distributed on the symbol period. The relative delays affect the distribution of the MAI, which would be simplified under the approximation of chip synchronism.

Previous work has addressed the error rate of optical CDMA receivers through the use of these simplifying approximations. A CDMA optical receiver using a post-photodetection (electrical) matched filter and Gold sequences was analyzed in [1]. The receiver decides for the data of the user of interest based on an observation of the optical intensity, rather than on a filtered point process driven by the intensity. This approximation is known as "perfect optical-to-electrical conversion," and ignores the quantized nature of light. Chip-synchronous transmission was also assumed, and the number of users was considered large enough to model the MAI as a Gaussian random variable. Finally, the dark current from the photodetector was ignored. A noncoherent, optical matched-filter CDMA receiver employing prime codes was analyzed in [2]. This receiver was not limited by the speed of electronic processing as in [1], since the matched-filter operation was performed optically. The authors assumed perfect optical-to-electrical conversion and Gaussian-distributed MAI. With these approximations the observation is a deterministic signal in Gaussian noise, and the authors demonstrated the superior performance of prime codes over Gold codes by a comparison of the signal-to-noise ratios.

Recently, a two-part paper explored the performance of an optical CDMA system using the optical matched-filter receiver [7,8]. The authors computed upper and lower bounds to the single-user error rate for those signature sequences whose periodic cross-correlations are limited to one chip of coincidence. The analysis ignores dark current and the quantized nature of light. With these approximations the observation is composed of the desired user's energy in the additive, aggregate interference, and the receiver compares this statistic to a threshold in order to decide on the transmitted information. Error rate bounds were obtained by considering bounds on the variance of the single-user interference. Since the bounds on the variance of each interferer were independent of the corresponding signature sequences, the authors avoided the need to compute all interference distributions. The upper bound to the variance was given by the chip-synchronous interference distribution, and the lower bound followed from an ideal set of codes in which the cross-correlations were strictly less than one chip of coincidence. The error rate bounds were compared as a function of the threshold, and differed by more than 2 orders of magnitude for most thresholds and numbers of transmitters of interest. Upper bounds were also computed for the error rate when the matched-filter receiver is preceded by an optical hard limiter. An ideal optical hard limiter is a nonlinear device which blocks the incident light for an input intensity below a fixed minimum value, and otherwise limits the output intensity to this minimum value. When the threshold of the hard limiter is set to the intensity of a single-user, the presence of the desired signal may be known exactly as before. In the absence of the desired signal, the intensity of the aggregate interference is reduced and bounded by that of a single user. In this way, an optical hard limiter reduces the error rate by clipping the aggregate interference intensity due to the overlap of two of more users' waveforms. Due to the nonlinear operation of the hard limiter, though, the output intensity is no longer the sum of independent signals, and an exact error analysis is complicated by this nonlinear transformation. The authors derived an upper bound to the error rate by overestimating the probability of interference in each non-zero chip of the desired signature sequence. This probability was bounded by randomly selecting one of the interferers in each non-zero chip to represent the interference, and the remaining users were permitted to interfere in the next chip position. Thus, each user could interfere repeatedly until he is chosen to represent the interference in a particular chip. This work compared the upper bounds to the error rate of the matched-filter CDMA receiver with and without an optical hard limiter, and showed that many more users may be accommodated with the incorporation of the optical hard limiter. In both upper bounds, the convolution of three of more iid uniform random variables on $[0, 1]$ was approximated by the Gaussian distribution.

An artifact of perfect optical-to-electrical conversion is "error-free" transmission, which occurs for a sufficiently small population of interferers. Under this approximation the observable is the total received energy in the symbol interval, which is additively composed of the desired user's energy and that of the interferers. Since the MAI has a maximum value which is proportional to the number of interferers, one may completely separate the ranges of the observable under each hypothesis if the number of interferers is small enough, and may specify an error-free threshold test. When the random nature of the photodetector output current is retained, this effect is lost.

In summary, recent analyses of noncoherent, optical CDMA receivers have relied on the approximations of perfect optical-to-electrical conversion, Gaussian-distributed MAI, and chip synchronism. The accuracy of these approximations is not known. How much do the approximations of Gaussian-distributed MAI and perfect optical-to-electrical conversion change the error rate? Does the approximation of chip synchronism yield an error rate that is close to the exact error rate? How small is the exact error rate when the "error-free" condition is satisfied?

In this work we derive the exact error rate for the noncoherent, optical matched-filter CDMA receiver, which decides for the data of a single user by comparing a photoelectron count to a threshold. The results adhere fully to the semi-classical model of light and do not depend on limit theorems for large user groups or strong received optical fields. The analysis is valid for arbitrary quantum efficiencies, binary signature sequences, random gain distributions, and dark currents, and is broad in application. In Section 2 we describe noncoherent, optical CDMA modulation and consider single-user demodulation based on a conditionally compound-Poisson observation. We derive the probability mass function (PMF) of the observation under each hypothesis in Section 3, and use them to determine the optimal threshold and minimum probability of error. What makes the analysis particularly interesting is the fact that while the formal representation of the PMF for a doubly-stochastic compound Poisson count is readily derived via conditioning [9], explicit forms are not common. Due to the particular nature of the MAI distribution, we are able to show that the PMF may be expressed as a straightforward summation. This expression is derived for independent and identically-distributed (iid) interferers having a distribution that includes the cases of user asynchronism and chip synchronism. It will be seen that the error rate expression is simplified when the distribution of the MAI is discrete. In Section 4 we take advantage of this fact to derive arbitrarily-tight bounds on the error rate which are considerably easier to compute than the exact error rate. In Section 5 we use the same bounding technique to derive arbitrarily-tight bounds on the error rate when the interferers' energies are not identical. In Section 6 we focus on the special case of prime codes, equal energies, and unity-gain photodetectors in order to compare the optimal threshold and minimum error rate to those obtained using the approximations discussed above. The approximation of perfect optical-to-electrical conversion yields poor estimates of the error rate and optimal threshold at moderate incident optical intensities and dark currents. Further, the combined approximations of perfect optical-to-electrical conversion and Gaussian-distributed MAI together yield an underestimate of the optimal threshold and an error rate that is neither an upper nor a lower bound. We also show that when prime sequences are employed, the chip-synchronous approximation leads to an overestimate of the error rate. The validity of these approximations for larger optical powers is also discussed in Section 6.

2. *Optical CDMA Model*

During each length-T symbol interval, the j^{th} transmitting laser is amplitude-modulated by the product of the data, which takes on values in $\{0,1\}$, and an assigned signature sequence. In this work, a signature sequence is a deterministic, $\{0,1\}$-valued, piecewise-constant function on $[0,T]$, and is specified by the values that it takes on the N equal-length subintervals (chips) of $[0,T]$. We define $P_j = P$ as the number of non-zero chips in the j^{th} signature sequence, b_{jn} as the transmitted symbol of the j^{th} user in the interval $(nT,(n+1)T]$, and $c_j(t)$ as a periodic replication of the j^{th} signature sequence. The transmitted complex scalar field from the j^{th} laser may be expressed as

$$r_j(t) = \sqrt{\frac{s\,N}{T}} c_j(t-\tau_j) b_{jn} e^{j(\nu_j[t-\tau_j]+a_j W_j(t-\tau_j)+\theta_j)}, \quad nT \leq t < (n+1)T, \tag{1}$$

where s is proportional to the optical energy per bit of the transmitting laser, ν_j denotes the optical carrier frequency of the j^{th} user, and θ_j is the phase offset of the j^{th} laser from the first laser. Thus we have assumed that all users transmit with identical signal energies. This is not the case in general, and in Section 5 we consider the more general case of unequal signal energies. The laser phase noise is represented by $a_j W_j(t)$, where a_j is related to the j^{th} transmitting laser linewidth, B_j, by $a_j = \sqrt{2\pi B_j}$. The relative delays $\{\tau_j\}_{j=2}^{K}$ are defined on $[0,T)$ with reference to the receiver of the first user. As there is no cooperation between the users, it is appropriate to model the relative delays as iid random variables that are uniformly distributed on the interval $[0,T)$. We shall assume that the symbol rate of each user is the same, the optical fields of the K users add in a noncoherent fashion, and that each single-user receiver acquires the timing of its transmitter's symbol epochs. With ideal transmission, (1) also represents the complex scalar field at the first receiver due to user j.

During the time interval $[0,T)$, the intensity of the total optical field at the receiver of the first user is

$$|r(t)|^2 = \frac{sN}{T} \sum_{j=1}^{K} b_{j-1} c_j(t-\tau_j) p_t(0,\tau_j) + b_{j0} c_j(t-\tau_j) p_t(\tau_j,T), \tag{2}$$

where $p_t(a,b)$ is a rectangular pulse of unit height with support $[a,b)$. We are interested in one-shot detection of the data b_{10} based on an observation of a photon count in $[0,T)$. The underlying photon point process is driven by a filtered version of $|r(t)|^2$, which depends on the data b_{10} only at times $\{t | c_1(t)=1,\ 0 \leq t < T\}$. This follows from (2) and the additional fact that $\tau_1 = 0$. By correlating the received point process in $[0,T)$ with $c_1(t)$, one may obtain the photon count during the support of c_1 and may then decide on the data b_{10} based on this count. This suboptimal processing scheme is the basis for the matched-filter CDMA receiver. Since the function $c_1(t)$ takes values on $\{0,1\}$, the correlation is easily achieved at low chip rates by an electro-optic modulator, which allows light to pass only when $c_1(t)=1$. A fiber optic tap delay line may be used to achieve the matched-filter operation at higher chip rates. This all-optical device uses the finite propagation velocity of light to achieve a relative delay between two optical signals by passing them through fibers of different lengths. The matched-filter CDMA receiver has been studied in several experiments [2,3] and will be the CDMA receiver analyzed in this work.

As seen in (1), this CDMA system employs a form of on-off modulation in which no light is transmitted for a "0", and the signature sequence is transmitted for a "1". This receiver has been analyzed previously using the point process model [10]. An optical CDMA system employing a different modulation scheme has also been analyzed using the point process model [11]. In that system each user is assigned two signature sequences, one for each symbol value. The receiver matched-filters the amplified photocurrent by the difference of the signature sequences of the desired user, and compares the output to a threshold.

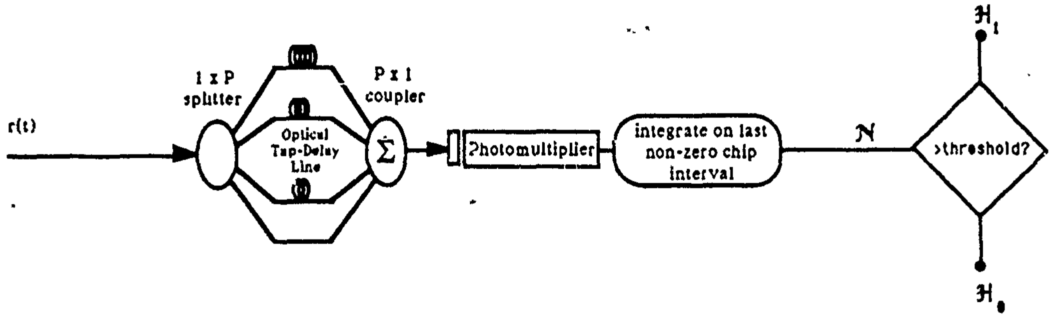


Figure 1. The optical, noncoherent matched-filter CDMA receiver

The matched-filter CDMA receiver to be analyzed in this paper is shown in Figure 1. The total received optical signal r(t) is coupled to a $1 \times P$ beam splitter. Each of the outputs of the splitter are identical copies of the input signal that are attenuated in intensity by P. The i^{th} tap delays the received field so that the optical signal during the i^{th} non-zero chip of the first signature sequence overlaps in time with the last non-zero chip in the undelayed signal. The delayed signals are noncoherently recombined, and the aggregate signal is incident on a photomultiplier screen. The photomultiplier emits a random number of secondary (output) electrons for each detected photon or thermoelectron, and the matched-filter receiver compares the secondary electron count during the last non-zero chip interval of the first signature sequence to a threshold in order to decide on the value of b_{10}. For the remainder of this work, we denote this secondary electron count by $\mathcal{N}$.

We shall employ a common photomultiplier model, in which the intensity of primary electrons is given by $\alpha|r(t)|^2 + \beta$, where α is proportional to the quantum efficiency of the photodetector, and β denotes the rate of primary electrons due to an independent dark current. The n^{th} primary electron yields a random number, g_n, of secondary (output) electrons, and the collection $\{g_n\}$ is modeled as an iid sequence, which is independent of the primary electron point process [12]. The common probability generating function of the random gains is denoted as $G(z) = \sum_{k=0}^{\infty} z^k \mathcal{P}[g_n = k]$. It follows that $\mathcal{N}$ is a conditionally compound-Poisson random variable given the integrated intensity, which we define as Λ,

$$\Lambda \triangleq \frac{1}{P}\int_0^T c_1(t)[\alpha|r(t)|^2 + \beta]\, dt = \alpha s\, b_{10} + d + \frac{\alpha s}{P}\sum_{j=2}^{K} b_{j-1}R_{j1}(\tau_j) + b_{j0}\hat{R}_{j1}(\tau_j), \tag{3}$$

and the conditional distribution of $\mathcal{N}$ depends only on $G(z)$ and Λ. Here $R_{j1}(\tau)$ and $\hat{R}_{j1}(\tau)$ are the normalized, partial cross-correlations

$$R_{j1}(\tau) \triangleq \frac{N}{T}\int_0^{\tau} c_j(t-\tau)c_1(t)dt$$

$$\hat{R}_{j1}(\tau) \triangleq \frac{N}{T}\int_{\tau}^{T} c_j(t-\tau)c_1(t)dt,$$

that represent the contributions to Λ by user j for the duration of b_{j-1} and b_{j0}, respectively. In (3) d represents the portion of the primary electron count mean due to thermoelectrons. Without loss of

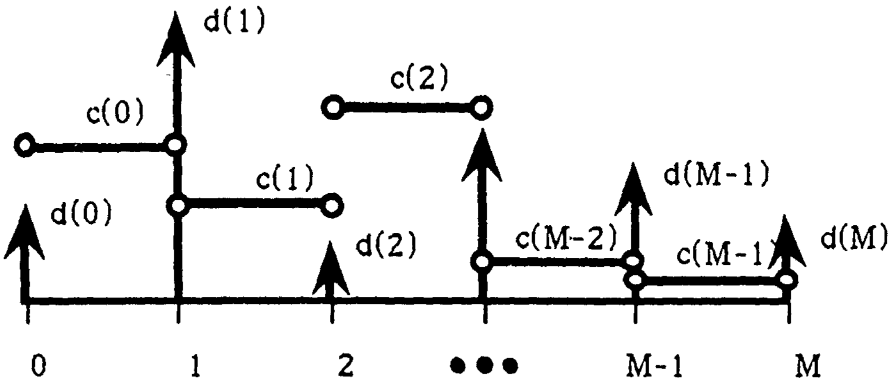


Figure 2. Probability density function of a γ-type random variable.

generality we set the quantum efficiency of the photodetector to unity, as this affects the distribution of $\mathcal{N}$ as an attenuation in intensity. For the same reason, we neglect the combining loss in the coupler. Further, we define $x \triangleq sb_{10} = 0$ under hypothesis $\mathcal{H}_0$ and $x = s$ under hypothesis $\mathcal{H}_1$

In summary, $\mathcal{N}$ is conditionally compound-Poisson-distributed with conditional primary electron mean Λ. This conditional mean has a distribution which is shifted to the right under $\mathcal{H}_1$, and the receiver decides for $\mathcal{H}_1$ if $\mathcal{N}$ exceeds some threshold. In general, a single-threshold test is not optimum, although sufficient conditions for the optimality of single-threshold detection have been determined [13-15]. An important outcome of this work is to determine the threshold which minimizes the error rate of the receiver. In other words, we shall find the minimum error rate detector among the class of detectors that compares the count $\mathcal{N}$ to a threshold. Since $\mathcal{N}$ is integer-valued, it is straightforward to find this threshold given the probability mass functions (PMFs) of $\mathcal{N}$ under each hypothesis. We find the PMFs of $\mathcal{N}$ in the next section.

3. Derivation of $\mathcal{P}[\mathcal{N} = n \mid x]$

In this section we derive the PMF of the secondary electron count at the integrator output for an arbitrary photomultiplier and for chip-synchronous or completely asynchronous transmission. This expression will be used in Section 4 to develop arbitrarily-tight, computationally-efficient bounds on the cumulative distribution function of $\mathcal{N}$, and in Section 6 we will use the probability mass functions of $\mathcal{N}$ in order to compare the exact error rates and optimal thresholds to those obtained using popular approximations.

In order to concisely describe the statistics of Λ, we define a γ-*type* random variable to be a mixed random variable having point masses at the integers $\{0, 1, \ldots M\}$ for some positive integer M, and constant, continuous portions between these integers. For fixed M, the class of γ-type random variables is parameterized by $2M$ parameters, each taking values on $[0, 1]$. If I is a γ-type random variable with $2M$ parameters, we define

$$d(i) \triangleq \mathcal{P}[I = i] \qquad i \in \{0, 1, \ldots, M\}$$

$$c(j)\, dv \triangleq \mathcal{P}[I \in [v, v + dv)] \qquad [v, v + dv) \subset (j, j + 1) \quad j \in \{0, 1, \ldots, M - 1\},$$

and we denote the distribution of I as $\{d(0), d(1), \ldots, d(M), c(0), \ldots, c(M - 1)\}$. Figure 2 illustrates the density of a γ-type random variable.

Each normalized, partial cross-correlation is a γ-type random variable with $2M$ parameters, where M is an upper bound on the set of total cross-correlations $\{R_{jk} + \hat{R}_{jk}\}$. This follows in part from the fact that each partial cross-correlation is piecewise-linear in the relative delay, since it is a convolution between piecewise-constant functions.† In fact, $R_{jk}(\tau)$ attains integer values at the chip-synchronous delays $\tau \in \{\frac{n}{N}T,\ n = 0, \ldots N-1\}$, and is piecewise-linear between these values. If the common distribution of the relative delays is discrete on $\{\frac{n}{N}T;\ n = 0, \ldots N-1\}$, then it is clear that $R_{jk}(\tau)$ is γ-type with $c(q) = 0\ \forall q$. It follows that under the assumption of chip synchronism, the distribution of the partial cross-correlations is γ-type. In the asynchronous case the relative delays are uniformly distributed on $[0, T]$, which combined with the piecewise linearity of $R_{jk}(\tau)$ also yields a γ-type distribution. In this case, point masses arise due to constant portions of the cross-correlations. Further, $c(q)$ is strictly positive if and only if the normalized, partial cross-correlation takes on both a value strictly greater than q and a value less than or equal to q on the set of chip-synchronous delays. From the piecewise-linearity of R_{jk}, it is also clear that the distribution of R_{jk} for user asynchronism may be computed from a knowledge of the function R_{jk} evaluated only at the chip synchronous delays. The $M = 1$, $\gamma-$type distribution has been used in the analysis of optical CDMA systems employing a subclass of OOCs [7]. This analysis, however, did not employ a point process model.

In this section we will derive the probability generating function (PGF) of $\mathcal{N}$ from its conditional compound-Poisson nature, and show that this z-transform has a particularly straightforward and explicit Maclaurin series expansion. The PMF is the collection of coefficients of this series, and may be explicitly represented, as follows. By conditioning on $\{x, (R_{21}, \hat{R}_{21}), \ldots (R_{K1}, \hat{R}_{K1})\}$, the count $\mathcal{N}$ has a compound Poisson distribution, whose PGF is given by (e.g., [16])

$$\mathrm{E}\left[z^{\mathcal{N}} \,\middle|\, x, (R_{21}, \hat{R}_{21}), \ldots (R_{K1}, \hat{R}_{K1})\right] = e^{(x+d)(G(z)-1)} \times \prod_{j=2}^{K} e^{\frac{s}{P}\{b_{j-1}R_{j1} + b_{j0}\hat{R}_{j1}\}(G(z)-1)}. \tag{4}$$

Due to the mutual independence of the pairs $\{(R_{j1}, \hat{R}_{j1})\}_{j=2}^{K}$, the closed form of the PGF $\mathrm{E}\left[z^{\mathcal{N}} \mid x\right]$ may be found by smoothing each factor in (4) individually. The j^{th} factor depends on the quantity $b_{j-1}R_{j1} + b_{j0}\hat{R}_{j1}$, which we describe as the j^{th} *interference mixture*. It is clear that the j^{th} interference mixture is also of γ-type with $2M$ parameters, and we denote its distribution as $\{D_j(0), D_j(1), \ldots, D_j(M), C_j(0), \ldots, C_j(M-1)\}$. Since the mixture coefficients $\{b_{j-1}, b_{j0}\}$ take values in the set $\{0, 1\}$, the distribution of the interference mixtures is computed easily from the marginals of the partial and total cross-correlations. With this notation, the closed-form expression of the power series of interest is

$$\begin{aligned} \mathrm{E}\left[z^{\mathcal{N}} \,\middle|\, x\right] = \mathrm{e}^{(G(z)-1)(x+d)} \times \prod_{j=2}^{K} \Bigg\{ & \sum_{q=0}^{M} D_j(q) \exp(q\frac{s}{P}(G(z)-1)) \\ & - \frac{P}{s} \frac{\mathrm{e}^{(G(z)-1)\frac{s}{P}} - 1}{1 - G(z)} \sum_{r=0}^{M-1} C_j(r) \exp(r\frac{s}{P}(G(z)-1)) \Bigg\}. \end{aligned} \tag{5}$$

We are interested in finding $\mathcal{P}[\mathcal{N} = n \mid x]$, the coefficient of z^n in the power series of (5) about $z = 0$. This power series is straightforward but unnecessarily general for most signature sequence sets of interest. For example, the number of parameters in the power series is reduced by a factor of $K-1$ by assuming that the distribution of the j^{th} interference mixture is independent of j - that is, the contribution of user j to the MAI is statistically indistinguishable from the other interferers. We have verified that this is a reasonable approximation when the signature sequences come from the prime codes, and will drop the subscript from the distribution of the interference mixtures for the sake of clarity. Also, the power

† We emphasize that in this work the signature sequences are deterministic and arbitrary. The randomness in the cross-correlations is due solely to the relative delays.

series of (5) is concisely written if we define $C(-1) \triangleq C(M) \triangleq 0$. With these simplifications, (5) becomes

$$\mathrm{E}\left[z^{\mathcal{N}} \,\middle|\, x\right] = e^{(G(z)-1)(x+d)} \times \left\{ \sum_{q=0}^{M} D(q) \exp(q \frac{s}{P}(G(z)-1)) - \frac{P}{s} \frac{1}{1-G(z)} \sum_{q=0}^{M} [C(q-1) - C(q)] e^{q\frac{s}{P}(G(z)-1)} \right\}^{K-1}. \qquad (6)$$

There are 2M terms inside of the braces. Letting n_q be the power of $D(q)$, and m_q be the power of $[C(q-1) - C(q)]$ in a term of a multinomial expansion of (6), we may express the PGF of $\mathcal{N}$ as

$$\mathrm{E}\left[z^{\mathcal{N}} \,\middle|\, x\right] = \sum \frac{(K-1)!}{\prod_{q=0}^{M} n_q! m_q!} \left\{ \prod_{q=0}^{M} D(q)^{n_q} \left[\frac{P}{s} \{C(q) - C(q-1)\} \right]^{m_q} \right\} \times \frac{\exp[(G(z)-1)\{x + d + \frac{s}{P} \sum_{l=0}^{M} l[n_l + m_l]\}]}{(1-G(z))^{\sum_{l=0}^{M} m_l}}, \qquad (7)$$

where the outer summation is over all the indices such that $\sum_{q=0}^{M} m_q + n_q = K - 1$. We find the PMF of $\mathcal{N}$ in the following way. Suppose that we knew the coefficients of the following power series

$$\sum_{n=0}^{\infty} \mathcal{W}(n, \alpha, \delta) z^n \triangleq \frac{e^{\alpha(G(z)-1)}}{(1-G(z))^{\delta}}, \quad \alpha \in \mathbb{R}_+, \ \delta \in \{0, 1, 2 \ldots\}. \qquad (8)$$

Recognizing the similarity of the last line of (7) with the right hand side of (8), we could express the PMF for $\mathcal{N}$ in terms of these coefficients as

$$\mathcal{P}[\mathcal{N} = n \mid x] = \sum \frac{(K-1)!}{\prod_{q=0}^{M} n_q! m_q!} \left\{ \prod_{q=0}^{M} D(q)^{n_q} \left[\frac{P}{s} \{C(q) - C(q-1)\} \right]^{m_q} \right\} \mathcal{W}\left(n, \{x + d + \frac{s}{P} \sum_{l=0}^{M} l[n_l + m_l]\}, \sum_{l=0}^{M} m_l \right), \qquad (9)$$

where the outer summation is over all m_q and n_q such that the $\sum_{q=0}^{M} m_q + n_q = K - 1$. In fact, $\mathcal{W}(n, \alpha, \delta)$ may be calculated by a linear recursion on the integers n and δ. This recursion for $\mathcal{W}$ is most easily seen by re-expressing the following identity using (8)

$$\frac{e^{\alpha(G(z)-1)}}{(1-G(z))^{\delta+1}} = G(z) \frac{e^{\alpha(G(z)-1)}}{(1-G(z))^{\delta+1}} + \frac{e^{\alpha(G(z)-1)}}{(1-G(z))^{\delta}}, \quad \delta \in \{0, 1, 2, \ldots\}.$$

We recall that $G(z)$ is the PGF of the photomultiplier gain distribution, $G(z) = \sum_{l=0}^{\infty} z^l \mathcal{P}[g = l]$. It follows from this substitution that for all $n, \delta \in \{0, 1, 2, \ldots\}$ the linear recursion for $\mathcal{W}$ is

$$(1 - \mathcal{P}[g = 0]) \mathcal{W}(n+1, \alpha, \delta+1) = \sum_{l=1}^{n+1} \mathcal{P}[g = l] \, \mathcal{W}(n+1-l, \alpha, \delta+1) + \mathcal{W}(n+1, \alpha, \delta) \qquad (10)$$

For most photomultiplier models $\mathcal{P}[g = 0] = 0$, which we will assume in the sequel. The initial conditions of this recursion are also easily extracted from the definition of $\mathcal{W}$,

$$\mathcal{W}(0, \alpha, \delta) = e^{-\alpha}, \quad \delta \in \{0, 1, 2, \ldots\} \qquad (11)$$

and

$$\mathcal{W}(n,\alpha,0) = \sum_{k=0}^{n} \frac{\alpha^k}{k!} e^{-\alpha} \ \mathcal{P}\left[\sum_{l=1}^{k} g_l = n\right] \ , \ \ n \in \{0,1,\ldots\}.$$

The linear recursion for $\mathcal{W}(j,\alpha,\delta)$ on n and δ permits fast, efficient computation for any arguments $n, \delta \geq 1$. Note that the second initial condition for this recursion depends on the probabilities $\mathcal{P}\left[\sum_{l=1}^{k} g_l = n\right]$, for $n, k \in \{0,1,2,\ldots\}$. These probabilities require iterated convolutions of the PMF of the random gain g_l, may be precomputed and stored for small n and k, and may be approximated accurately for large n and k. It is easy to show that $\mathcal{P}\left[\sum_{l=1}^{k} g_l = n\right]$ has an explicit form for random gains that are shifted-Poisson-distributed in addition to the unity-gain case.

It is helpful at this point to look at (9) in the special case of one interferer and unity-gain photodetection. In this case (9) becomes

$$\begin{aligned}\mathcal{P}[\mathcal{N} = n \mid x] = &\sum_{j=0}^{M} D(j)\mathcal{P}\left[\Pi(d + x + j\frac{s}{P}) = n\right] \\ &+ \sum_{j=0}^{M-1} C(j)\left\{\mathcal{P}\left[\Pi(d + x + j\frac{s}{P}) \leq n\right] - \mathcal{P}\left[\Pi(d + x + (j+1)\frac{s}{P}) \leq n\right]\right\}\end{aligned}$$

where $\Pi(m)$ is a Poisson random variable with mean m. From this equation, we see that if the interference takes on a value corresponding to a point mass, then the contribution to the count PMF is a Poisson PMF, as expected. If the interference takes on a value between the point masses, then the contribution to the PMF of $\mathcal{N}$ is the difference between two Poisson cumulative distribution functions (CDFs). This latter fact is due to the piecewise constant nature of the continuous portion of the interference distribution.

We have shown that the PMF of $\mathcal{N}$ may be expressed as a finite summation of terms involving the function $\mathcal{W}$. We have also demonstrated that $\mathcal{W}$ may be computed by a linear recursion on its integer arguments. These expressions are valid when the distribution of the normalized, partial cross-correlations is γ-type (Figure 2) such as in the cases of user asynchronism and chip-synchronism. It should be pointed out that the number of terms in the PMF summation (9) is exponential in the number of interferers, and may be prohibitively large for large K. In the next section we show that by modifying the distribution of the interference mixtures, we may use a special case of (9) to express arbitrarily tight bounds on the PMF as a summation that is bilinear in the number of interferers $K-1$ and the parameter M.

4. *Arbitrarily Tight Bounds on* $\mathcal{P}[\mathcal{N} \leq n \mid x]$

Computationally-efficient bounds must reduce the complexity of (9) in both the multinomial summation and the computation of $\mathcal{W}$, while controlling the loss of accuracy by a parameter of our selection. In this section we show that by quantizing the interference mixtures, we achieve all three objectives. The intuition behind the bounds is demonstrated in the following special case. Suppose that each interference mixture has a discrete distribution on the set $\{0, \frac{1}{Q}, \frac{2}{Q}, \ldots, M\}$ containing $QM + 1$ elements. In this case the conditional mean Λ takes on $(K-1)QM + 1$ possible values. When Λ has a discrete distribution, $C(j) = 0 \ \forall j$, and the conditional CDF of $\mathcal{N}$ given x and all partial cross-correlations is given by (9) as

$$\mathcal{P}\left[\mathcal{N} \leq n \,\middle|\, x, (R_{21}, \hat{R}_{21}) \ldots (R_{K1}, \hat{R}_{K1})\right] = \sum_{t=0}^{n} \mathcal{W}(t, \Lambda, 0). \qquad (12)$$

Since the right-hand side depends only on an initial condition for $\mathcal{W}$, the assumption of a discrete distribution on Λ eliminates the need to compute the linear recursion for W. Also, computation is reduced substantially by noting that the conditional CDF of $\mathcal{N}$ depends on $\{R$
the conditional mean $\Lambda = x + d + \frac{s}{P}\sum_{j=2}^{K} b_{j-1}R_{j1} + b_{j0}\hat{R}_{j1}$. This is not true for the exact interference distribution, in which case the conditional CDF is also dependent on $\delta = \sum_{j=2}^{K} m_j$, the total number of interference mixtures that take on non-integer values. The effect of δ on the exact PMF may be seen in (9). This observation allows us to uncondition (12) over the distribution of the $(K-1)QM+1$-valued conditional primary electron mean Λ, rather than the joint distribution of the interference mixtures, which would require $(2M+1)^{K-1}$ terms. Quite simply, a discrete distribution for the interference mixtures yields a CDF for $\mathcal{N}$ that is linear rather than exponential in the number of interferers.

But how do we obtain arbitrarily-tight bounds on the exact conditional error rate $\mathcal{P}[\mathcal{N} \leq n \mid x]$ that use a discrete interference mixture distribution? Suppose we quantize each term of the interference mixture $b_{j-1}R_{j1} + b_{j0}\hat{R}_{j1}$ with a step size of $\frac{1}{Q}$, $Q \in \{1, 2, \ldots\}$, and round-up or round-down to form bounds on the interference mixtures. That is, we form Λ_l, Λ_u from (3)

$$\Lambda_l = x + d + \frac{s}{P}\sum_{j=2}^{K} b_{j-1}\frac{1}{Q}\lfloor QR_{j1}\rfloor + b_{j0}\frac{1}{Q}\lfloor Q\hat{R}_{j1}\rfloor$$

and

$$\Lambda_u = x + d + \frac{s}{P}\sum_{j=2}^{K} b_{j-1}\frac{1}{Q}\lceil QR_{j1}\rceil + b_{j0}\frac{1}{Q}\lceil Q\hat{R}_{j1}\rceil,$$

where $\lfloor R\rfloor$ ($\lceil R\rceil$) is the greatest (least) integer function of R. Although it is obvious that $\Lambda_l \leq \Lambda \leq \Lambda_u$, it is not clear that we may form bounds on the secondary electron count CDF by substituting Λ_l and Λ_u for Λ. A subtle point is raised by considering the form of $\mathcal{N}$

$$\mathcal{N}(\Lambda) = \sum_{p=1}^{\Pi(\Lambda)} g_p \tag{13}$$

where $\Pi(\Lambda)$ is a conditionally-Poisson count with conditional mean Λ. Since the random gains g_p are non-negative, $\mathcal{N}$ increases with the primary electron count, Π. It is easy to show that $\Lambda_l < \Lambda$ implies $\mathcal{P}[\Pi(\Lambda) \leq n] < \mathcal{P}[\Pi(\Lambda_l) \leq n]$ for all n [17], yet it is not clear from (13) that corresponding bounds on the CDF of $\mathcal{N}$ follow from this fact. In the lemma below we show that we may achieve bounds on $\mathcal{P}[\mathcal{N} \leq n \mid x]$ by using the distributions of Λ_l and Λ_u. This is shown by first considering the case when Λ is deterministic.

Lemma 1. Let $\Pi(a)$ be a Poisson random variable with mean a, and let $\mathcal{N}(a) = \sum_{k=1}^{\Pi(a)} g_k$, where $\{g_k\}$ is a collection of nonnegative i.i.d. random variables that is independent of $\Pi(a)$. Let $0 < a' \leq a$. Then

$$\mathcal{P}[\mathcal{N}(a) \leq n] \leq \mathcal{P}[\mathcal{N}(a') \leq n] \ , \ n \geq 0.$$

Proof.

$$
\begin{aligned}
\mathcal{P}[\mathcal{N}(a) \le n] &= \sum_{k=0}^{\infty} \mathcal{P}[\Pi(a) = k]\, \mathcal{P}[g_1 + \ldots + g_k \le n] \\
&= \sum_{k=0}^{\infty} \{\mathcal{P}[\Pi(a) \le k] - \mathcal{P}[\Pi(a) \le k-1]\}\, \mathcal{P}[g_1 + \ldots + g_k \le n] \\
&= \sum_{k=0}^{\infty} \mathcal{P}[\Pi(a) \le k]\, \{\mathcal{P}[g_1 + \ldots + g_k \le n] - \mathcal{P}[g_1 + \ldots + g_{k+1} \le n]\} \\
&\le \sum_{k=0}^{\infty} \mathcal{P}[\Pi(a') \le k]\, \{\mathcal{P}[g_1 + \ldots + g_k \le n] - \mathcal{P}[g_1 + \ldots + g_{k+1} \le n]\} \\
&= \mathcal{P}[\mathcal{N}(a') \le n]
\end{aligned}
$$

The inequality follows from the fact that the Poisson CDF is a decreasing function of the mean. This lemma extends easily to the case when $\Pi(\Lambda)$ is conditionally Poisson with conditional mean Λ. By conditioning on Λ, we have

$$\mathcal{P}[\mathcal{N}(\Lambda) \le n \,|\, \Lambda] \le \mathcal{P}[\mathcal{N}(\Lambda_l) \le n \,|\, \Lambda]$$

from Lemma 1 since both Λ and Λ_l are known under this conditioning. The desired result follows by smoothing with respect to the distribution of Λ. Lemma 1 allows us to compute an upper bound of the CDF of $\mathcal{N}$ under each hypothesis by averaging the right hand side of

$$\mathcal{P}[\mathcal{N}(\Lambda) \le n \,|\, x, \Lambda, \Lambda_l] \le \mathcal{P}[\mathcal{N}(\Lambda_l) \le n \,|\, x, \Lambda, \Lambda_l]$$

over the $(K-1)QM + 1$-valued distribution of Λ_l. The tightness and complexity of the upper (and lower) bounds are controlled by the quantization step size $\frac{1}{Q}$. It is obvious by comparing the number of terms per point of the CDF to those of the bounds that even for moderate numbers of transmitters and fine quantization; the bounds require significantly less computation.

5. *Arbitrarily-Tight Error Rate Bounds for Unequal Energies*

Using Lemma 1, we may form bounds on the count CDF under each hypothesis by developing bounds on the underlying conditional mean. In the case of equal energies, we may achieve this by uniformly quantizing each interference mixture, and as the quantization step size decreases, the accuracy of the bounds improves. In this section we show that the same technique may be applied to the case of unequal energies.

We may express the conditional mean as

$$\Lambda = x + d + \frac{1}{P} \sum_{j=2}^{K} s_j [b_{j-1} R_{j1} + b_{j0} \hat{R}_{j1}],$$

where s_j is the energy per "1" for user j and $x = s_1 b_{10}$. In this case we define Λ_l, Λ_u as

$$\Lambda_l = x + d + \sum_{j=2}^{K} b_{j-1} \frac{1}{Q} \lfloor Q \frac{s_j}{P} R_{j1} \rfloor + b_{j0} \frac{1}{Q} \lfloor Q \frac{s_j}{P} \hat{R}_{j1} \rfloor$$

$$\Lambda_u = x + d + \sum_{j=2}^{K} b_{j-1}\frac{1}{Q}\lceil Q\frac{s_j}{P}R_{j1}\rceil + b_{j0}\frac{1}{Q}\lceil Q\frac{s_j}{P}\hat{R}_{j1}\rceil.$$

Note that the quantity $\frac{s_j}{P}$ appears inside each quantizer. Unlike the equal-energy case, we cannot form Λ_l and Λ_u by quantizing the partial cross-correlations directly. Instead, each must be scaled by the appropriate energy s_j. However, it is clear that $\Lambda_l < \Lambda < \Lambda_u$, and we may form bounds on the count CDF using quantization and Lemma 1. The quantized mean may take on at most $(MQ\max_j \frac{s_j}{P} + 1)^{K-1} - (K-1)$ values, where M is the maximum for the common interference mixture. The distributions of Λ_l and Λ_u are obviously discrete, and may be computed exactly. More importantly, we may uncondition (12) by the distribution of Λ_u to obtain a lower bound to the count CDF under each hypothesis. The bounds on the CDF of $\mathcal{N}$ are most easily seen from (9) with the following abuse of notation. Let $M = Q\ \max(\Lambda_u - (x+d))$, and let $D(j) = \mathcal{P}\left[\Lambda_u = x + d + \frac{j}{Q}\right]$ denote the distribution of Λ_u in (9). Since $\max(\Lambda_u - (x+d))$ is a multiple of $1/Q$ due to quantization, M is an integer. The idea is to consider Λ_u as the quantized interference due to *one* user having an interference mixture distribution $\{D(0), \ldots D(M)\}$. To complete this analogy, we must set $K - 1 = 1$ in (9). The corresponding CDF bound is

$$\mathcal{P}[\mathcal{N} \le n \mid x] \ge \sum_{i=0}^{n}\sum_{j=0}^{M} D(j)\mathcal{W}(i, x + d + \frac{j}{Q}, 0)$$

Note that we have set $s = 1$ in (9), since the interference energies are incorporated in the quantization. An upper bound to $\mathcal{P}[\mathcal{N} \le n \mid x]$ follows from Λ_l in the same way. It is helpful to note that for unity-gain photodetection the above bound becomes

$$\mathcal{P}[\mathcal{N} \le n \mid x] \ge \sum_{j=0}^{M} D(j)\mathcal{P}\left[\Pi(x + d + \frac{j}{Q}) \le n\right]$$

where $\Pi(m)$ is a Poisson random variable with mean m. This follows from the conditional Poisson nature of $\mathcal{N}$, when driven by a conditional mean with a discrete distribution.

6. *Example: Equal Energies, Prime Sequences and PIN Photodiodes*

In order to compare the exact error rate to the approximations discussed earlier, we must first compute the γ-type distribution $(D(0), \ldots, D(M), C(0), \ldots, C(M-1))$, which is used in (9). This distribution may be computed once the signature sequence set and the distribution of the relative delays are specified. In this section we focus on the set of prime sequences [18] in the user-asynchronous and chip-synchronous cases. We shall also assume equal energies for all users.

Since the normalized cross-correlations of prime sequences are bounded above by $M = 2$, we must compute $\{D(0), D(1), D(2), C(0), C(1)\}$ for the chip-synchronous and asynchronous cases [18]. For the prime sequences from GF(31), we have found that the average distributions for the interference mixtures are given to two significant digits as

	D(0)	D(1)	D(2)	C(0)	C(1)
chip-synchronous users ⇒	.57	.36	.07	.00	.00
asynchronous users ⇒	.44	.22	.01	.24	.09

Table 1. Average distributions of the interference mixtures.

We have verified that the MAI for prime sequences is well-modeled by a sum of iid random variables. In particular, the mean, variance, and third central moment using the average distribution for each interferer were identical to the exact MAI moments, while the fourth central moments differed by less than .004% for 29 interferers. Further, the average distribution of the GF(31) prime sequences did not differ significantly from those of the GF(11) and GF(17) prime sequences. For this reason we shall use the distribution in Table 1 for all calculations. It should be pointed out that bounds may be achieved by considering best and worst case (common) interference distributions, as was done in [8].

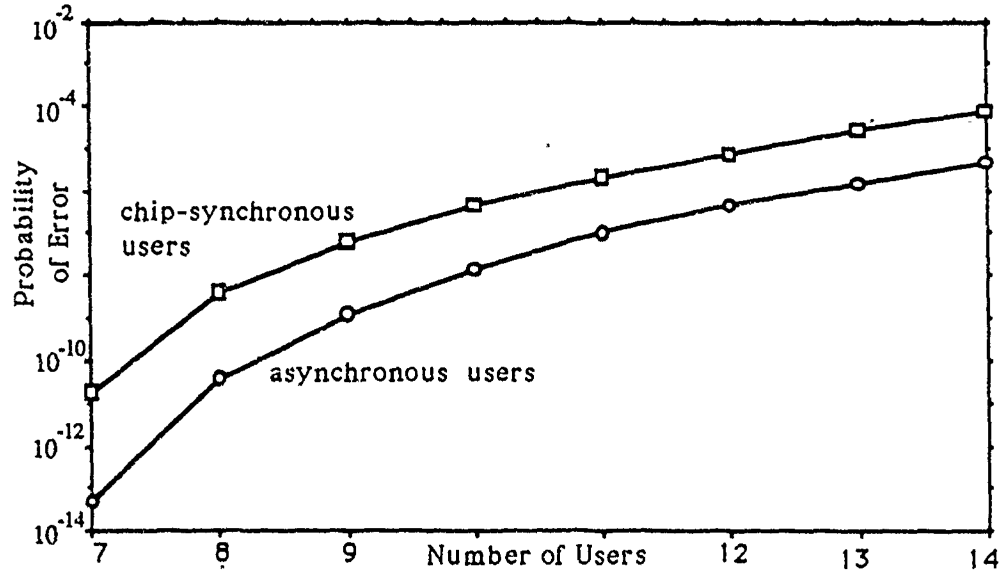


Figure 3. Comparison of the Minimum Error Rates For Asynchronous Users and the Chip-Synchronous Approximation

In Figure 3 we have plotted the minimum error probability of the matched-filter CDMA receiver for the chip-synchronous approximation and for completely asynchronous transmission. We have used the weight 17 and length 289 prime sequences from GF(17), an optical energy per user of s=1000 photons per bit, and a dark current mean of d=50 thermoelectrons per bit. For a single-user transmission rate of $\mathcal{R}$ Gigabits per second, these numbers correspond to a peak received single-user power of $\mathcal{R}$ μW and a photodetector dark current of approximately $10 \cdot \mathcal{R}$ nA. From Figure 3 we see that the chip-synchronous approximation upper-bounds the error rate in the asynchronous case by at least one order of magnitude. The error rates are ordered in this way due exclusively to the differences of the distributions of the interference mixtures. From Table 1, it may be shown that the means of the interference mixtures are identical in both cases, while the ordering of the variances coincides with that of the error rates. Thus the MAI has identical means under these distributions, and second moments whose ordering coincides with that of the error rates. In the unity-gain case it is easy to show that $E[\mathcal{N} \mid x] = \bar{\Lambda}$, and $Var(\mathcal{N}|x) = \bar{\Lambda} - (\bar{\Lambda})^2 + \overline{\Lambda^2}$, which implies that under each hypothesis the mean of $\mathcal{N}$ is unchanged by the approximation of chip synchronism, yet the variance of $\mathcal{N}$ increases as we proceed from complete asynchronism to chip synchronism. From the ordering of the minimum error rate curves in Figure 3, we see that an increase in the variance of $\mathcal{N}$ under each hypothesis results in an increased error rate as the conditional means of $\mathcal{N}$ are fixed.

In [7] it was shown by example that the variance of the interference mixture increases under chip synchronism for optical orthogonal codes that are bounded by one chip of interference. We will now show that an interference mixture $b_{j-1}R_{j1} + b_{j0}\hat{R}_{j1}$ has a larger variance under chip synchronism than asynchronism for arbitrary, deterministic signature sequences. Due to the independence of the users, this implies that the MAI has a larger variance under chip synchronism. We will prove this in two steps: we shall show first that each cross-correlation $R_{j1}, \hat{R}_{j1}$, and $R_{j1} + \hat{R}_{j1}$ has a larger variance under chip-synchronism, and then we shall show that the same is true for the interference mixture.

Lemma 2. The variance of the MAI is greater under chip synchronism than complete asynchronism.

Proof. Let $R(\tau)$ denote any of the above cross-correlations. Then the distribution of $R(\tau)$ is γ-type, and $R(\tau)$ has an upper bound of, say, M. $R(\tau)$ is a piecewise linear, continuous function of the relative delay $\tau \in [0,T]$. In fact, $R(\tau)$ is linear between its chip-synchronous values. Due to the uniform distribution on the delay, and the piecewise linearity of $R(\tau)$, it is easy to show that the mean of $R(\tau)$ is the same under chip synchronism or asynchronism. (The same is true for the interference mixture by similar arguments.) Thus, it is sufficient to show that the second moment of $R(\tau)$ is greater under chip synchronism than complete asynchronism.

The distribution of $R(\tau)$ under complete asynchronism has the form $(d(0), \ldots d(M), c(0), \ldots, c(M-1))$, where $c(j)$ represents a piecewise constant portion between j and $j+1$. Equivalently, this distribution may be described by $(d(0), \ldots, d(M), \{u_{mk}\}$, for $m, k = 0, \ldots M$ and $m < k)$, where $\frac{u_{mk}}{|m-k|}$ is the height of a square pulse on the interval (m,k). These square pulses overlap, and the sum of all $\frac{u_{mk}}{|m-k|}$ such that $(j, j+1) \subset (m,k)$ is equal to $c(j)$. The value u_{mk} is equal to the fraction of consecutive values of $\{m,k\}$ or $\{k,m\}$ in the sequence of chip synchronous values of $R(\tau)$. We introduce this decomposition for the following two reasons. First, the second moment of $R(\tau)$ is linear in each of the u_{mk}. Second, the probability mass on the interval (m,k) represented by u_{mk} under user asynchronism vanishes to the endpoints m and k under chip synchronism in such a way as the center of mass is conserved on the interval (m,k). The latter fact is the reason that the means of $R(\tau)$ coincide under chip synchronism and complete asynchronism. Because of these two facts, it is sufficient to show that under the constraint of a constant mean, the uniform distribution on (m,k) has a smaller second moment than the discrete distribution on $\{m,k\}$. This condition is easy to show, and it follows that the variance of any cross-correlation $R_{j1}, \hat{R}_{j1}$, or $R_{j1} + \hat{R}_{j1}$ between deterministic sequences is greater under chip synchronism than asynchronism.

Now we would like to show that the interference mixture $b_{j-1}R_{j1} + b_{j0}\hat{R}_{j1}$ has a higher variance under chip synchronism. Since the means of this random variable coincide under chip synchronism and complete asynchronism, it is sufficient to show that the second moment of the interference mixture is larger under chip synchronism. But this follows from the same fact for the cross-correlations, since the distribution of the interference mixture is a convex combination of the marginals of $R_{j1}, \hat{R}_{j1}$, and $R_{j1} + \hat{R}_{j1}$.

Since the variance of each interference mixture increases under chip synchronism, the same is true for the MAI.■

Direct-detection communication systems often require large received optical energies to achieve an acceptable error rate. Therefore, we are interested in the asymptotic distribution of $\mathcal{N}$ as s grows without bound. In the simplest case when Λ is deterministic ($\mathcal{N}$ is compound-Poisson) it is well-known that a normalized version of $\mathcal{N}$ converges in distribution to a Gaussian random variable. The asymptotic distribution has also been established for random Λ and unity-gain photodetection [19,20]. This result requires that $\bar{\Lambda} \to \infty$ and $\sigma_\Lambda > 0$, and shows that the asymptotic distribution depends on the limit of $\rho \triangleq \frac{\sigma_\Lambda^2}{\bar{\Lambda}}$. If $\lim \rho = 0$, then the normalized count converges in distribution to a standard Gaussian random variable. If $\lim \rho = \infty$, then the normalized count converges in distribution to $\tilde{\Lambda} \triangleq \lim \frac{\Lambda - \bar{\Lambda}}{\sigma_\Lambda}$.

Finally, if $0 < \lim \rho < \infty$, then the normalized count converges in distribution to an independent mixture of $\tilde{\Lambda}$ and a standard Gaussian. This result has also been generalized to include the case of photomultiplication [21]. We state this result below.

Lemma 3. Let Π be a conditionally Poisson random variable with mean Λ when conditioned on Λ, and let $\mathcal{N} = \sum_{k=0}^{\Pi} g_k$, where $\{g_k\}$ are iid positive random variables with finite mean and variance. Suppose that $\{g_k\}$ are independent of Λ, $\sigma_\Lambda > 0$, and define $\rho \triangleq \frac{\sigma_\Lambda^2}{\bar{\Lambda}}$. Then if $\bar{\Lambda} \rightarrow \infty$ the normalized random variable $\tilde{\mathcal{N}} \triangleq \frac{\mathcal{N}-\bar{\mathcal{N}}}{\sigma_{\mathcal{N}}}$ converges in distribution as follows. If $\lim \rho = 0$, then $\tilde{\mathcal{N}}$, converges to a standard Gaussian random variable. If $\lim \rho = \infty$, then $\tilde{\mathcal{N}}$, then converges to $\tilde{\Lambda}$. Finally, if $\lim \rho$ is finite and non-zero, then $\tilde{\mathcal{N}}$ converges to a mixture of the two limiting cases above.∎

Lemma 3 is most easily shown by taking the limit of the characteristic function of $\tilde{\mathcal{N}}$. In the matched-filter CDMA receiver, Λ is proportional to s, and $\bar{\Lambda}, \rho \rightarrow \infty$. Therefore, the normalized count converges in distribution to the scaled conditional mean $\tilde{\Lambda}$ as $s \rightarrow \infty$. This asymptotic result is more commonly known as "perfect optical-to-electrical conversion". It is important to note that $\tilde{\Lambda}$ is not a Gaussian random variable, as it is composed of a finite number of bounded, mixed random variables. It follows from Lemma 3 that the normalized count will not converge to a Gaussian random variable as $s \rightarrow \infty$, in contrast to the case when Λ is deterministic. In fact, it will be shown in the numerical results that the observed count $\mathcal{N}$ is poorly approximated by a signal in additive Gaussian noise.

In Figure 4 we compare the exact minimum error rate of the CDMA matched-filter receiver to those obtained with simplifying approximations. For these calculations we have chosen the signature sequences from the GF(11) prime codes, which have a length of 121 chips and a weight of 11 chips. We have also assumed that the average number of thermoelectrons is much less than the average number of photoelectrons. As expected, the exact minimum error rate increases with the number of users, and is a decreasing function of the single-user power. We see that the error rate curves seem to converge to that predicted by Lemma 3 as s increases. However, the asymptotic curve is a lower bound to the exact error rate by an order of magnitude for optical powers less than 10,000 photons per bit. Also note that each exact error rate curve (constant s) approaches the asymptotic curve corresponding to Gaussian-distributed MAI as the number of users, K, increases. This fact may also be justified by Lemma 3. However, in this case both $\bar{\Lambda}$ and σ_Λ^2 are proportional to $(K-1)$, and the limiting value of ρ is finite. We conclude from Lemma 3 that the asymptotic distribution of $\mathcal{N}$ for large K is a mixture of the random variable $\tilde{\Lambda}$ and an independent Gaussian random variable. In addition, $\tilde{\Lambda}$ converges in law to the Gaussian distribution by the central limit theorem as the number of users increases. Therefore, for fixed single-user power and increasing number of users, $\tilde{\mathcal{N}}$ converges weakly to a Gaussian random variable. This fact is illustrated in Figure 4, where the asymptotic error rate curve approaches the error rate curve for the approximation of Gaussian-distributed MAI. However, we note that the error rate is unacceptably high in the region in which the Gaussian-distributed MAI approximation is tight. It is also evident from Figure 4 that, in general, the Gaussian-distributed MAI approximation yields a poor estimate of the system performance.

Using the numerical results in Figure 4 we may also address the "error-free" condition, an artifact of perfect optical-to-electrical conversion. Under this approximation, the observation is additively composed of the desired user's energy and the MAI. Since the desired signal is on-off keying, the "error-free" condition exists when the signal peak is greater than the maximum of the MAI. Since the prime codes have a cross-correlation bound of 2, this condition occurs when the number of interferers is less than half the weight of the sequences. Therefore, the "error-free" condition should occur for less than 6 users in Figure 4. The exact error rate curves in this figure indicate that "error-free" performance is approximated for incident optical energies exceeding 10,000 photons per bit - the error rate for K=6

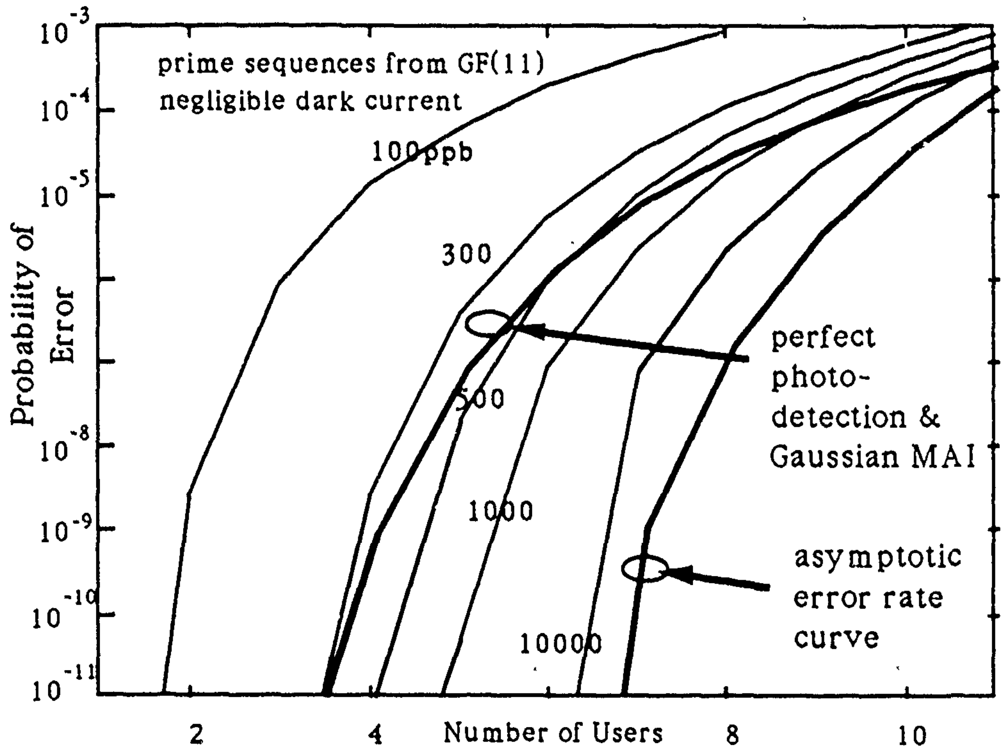


Figure 4. Exact error rate curves vs. number of users for various optical energies. All users have equal signal energies. Also shown are the curves corresponding to the high energy limit, and Gaussian-distributed MAI.

at this energy is roughly 10^{-14}. At optical energies less than 10,000 photons per bit, the "error-free" phenomenon is not observed.

An important byproduct of this analysis is the optimal threshold function for the matched-filter CDMA receiver. In Figure 5 we have plotted the optimal thresholds for those error rate curves plotted in Figure 4. Each threshold function is normalized by the respective signal energy per bit. As the incident optical energy per bit increases, the normalized optimal threshold increases to unity, which is the curve corresponding to the asymptotically optimal test. Note that the combined approximations of Gaussian-distributed MAI and perfect optical-to-electrical conversion yield a threshold that significantly underestimates the exact optimal threshold for moderate and large received optical energies. For this region the high-energy test (using the exact MAI distribution) yields a more accurate estimate of the optimal threshold. This fact further illustrates that the observable is not well-modeled as a Gaussian random variable for any optical power. Optimal thresholds for large incident optical energies are not plotted for the "error-free" region because they could not be reliably determined due to the vanishing error rate.

7. Conclusions

In this paper we have presented a semi-classical analysis of the error rate for a noncoherent,

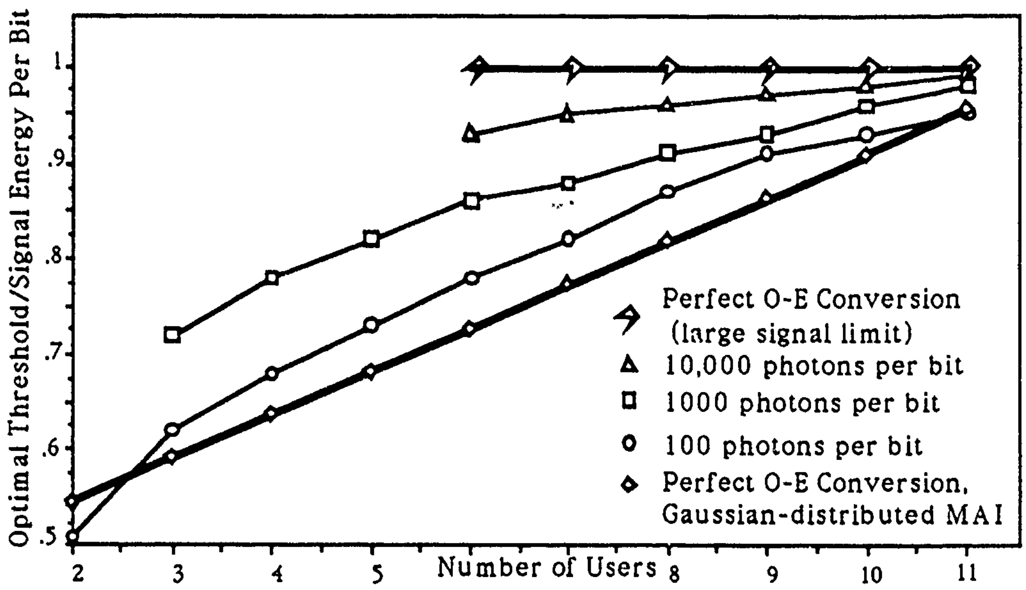


Figure 5. Optimal thresholds for the matched-filter CDMA receiver

matched-filter CDMA receiver in the optical channel. The error rate expression is valid for a common photomultiplier model, arbitrary signature sequences and equal energies among the users. In this paper we also developed arbitrarily-tight bounds on the error rate for unequal energies. In the case when the signal energies coincide, these bounding expressions are considerably easier to compute than the exact error rate. The exact error rates and optimal thresholds were compared to those obtained through various approximations for the special case of prime sequences and unity-gain photodiodes, and the accuracy of these approximations was addressed for various received optical energies and numbers of transmitters. It was demonstrated that the chip-synchronous approximation yielded minimum error rates that upper-bounded the exact error rates, that the approximation of perfect optical-to-electrical conversion was accurate only when the single-user optical energy per bit was much larger than 10000 photons per bit, and that the combined approximations of perfect optical-to-electrical conversion and Gaussian-distributed MAI were appropriate only for numbers of transmitters that yielded unacceptable error rates for a moderate number of chips per bit.

8. References

[1] S. Tamura, S. Nakano, K.Okazaki, *"Optical Code-Multiplex Transmission by Gold Sequences"* JOURNAL OF LIGHTWAVE TECHNOLOGY LT-3, No1, *February, 1985, pp.121-127*

[2] P.R. Prucnal, M.A. Santoro, T.R. Fan, *"Spread Spectrum Fiber-Optic Local Area Network Using Optical Processing"*, JOURNAL OF LIGHTWAVE TECHNOLOGY Vol. LT-4, No, 5, *May,1986, pp.547-554*

[3] J.A. Salehi, C.A. Brackett, *"Fundamental Principles of Fiber Optic Code Division Multiple Access (FO-CDMA)"*, INTERNATIONAL CONFERENCE ON COMMUNICATIONS, *1987, pp.1601-1609*

[4] E. Hoverston, *"Optical Communication Theory,"* in *Laser Handbook,* EDS. F.T. ARECCHI AND E.O. SCHULTZ, NORTH HOLLAND, AMSTERDAM, *1972, pp.1805-1862*

[5] L. Mandel, *"Fluctuations of Light Beams,"* in *Progress in Optics,* **Vol. II,** *ed. E.Wolf, Wiley, New York, 1963*

[6] L. Mandel, *"Fluctuations of Photon Beams and Their Correlations,"* PROC. PHYS. SOC. (LONDON), **72-1,** *1958, pp.1037-1048*

[7] J.A. Salehi, *"Code Division Multiple-Access Techniques in Optical Fiber Networks- Part I: Fundamental Principles",* IEEE TRANSACTIONS ON COMMUNICATIONS, **Vol. 37, No.8,** *August, 1989, pp.824-833*

[8] J.A. Salehi, *"Code Division Multiple-Access Techniques in Optical Fiber Networks- Part II: Systems Performance Analysis",* IEEE TRANSACTIONS ON COMMUNICATIONS, **Vol. 37,No.8,** *August, 1989, pp.834-842*

[9] D.L.Snyder, *Random Point Processes,* J. WILEY & SON, *1975*

[10] D. Brady, S.Verdú, *"An Exact Analysis of the Optical CDMA Noncoherent Receiver",* PROCEEDINGS OF THE CONFERENCE ON INFORMATION SCIENCES AND SYSTEMS, *Johns Hopkins University, March,1989 pp. 327-332*

[11] A. Lam, A. Hussain, *"Performance Analysis of Code Division Multiple Access Optical Communications with Avalanche Photo Detectors",* PROCEEDINGS ON THE CONFERENCE ON INFORMATION SCIENCES AND SYSTEMS, *Johns Hopkins University, March, 1989, pp.514-519*

[12] M.C. Teich, K. Matsuo, B.E.A. Saleh, *"Counting Distributions and Error Probabilities for Optical Receivers Incorporating Superlattice Avalanche Photodiodes",* ELECTRON DEVICES, **ED-33** *Oct, 1986, pp.1475-1488*

[13] P.R. Prucnal, *"Single-Threshold Processing for Orthogonal Likelihood-Ratio Detection with Application to FSK Fiber-Optic Communications"* IEEE TRANSACTIONS ON COMMUNICATIONS, **Vol. COM-29, No. 5,** *May, 1981, pp.743-749*

[14] P.R. Prucnal, M.C.Teich, *"Single-threshold detection of a random signal in noise with multiple independent observations. 1: Discrete case with applications to optical communications",* APPLIED OPTICS, **Vol. 17, No. 22,** *Nov. 1978, pp.3576-3583*

[15] D. Brady, *"Analyses of Optical, Direct-Detection Communication Systems Using a Point Process Model," PhD. Dissertation, Princeton University, May, 1990*

[16] D.L.Snyder, *Random Point Processes,* J. WILEY & SON, *1975, p.132*

[17] D. Brady, S. Verdú, *"Performance Analysis of an Asymptotically Quantum-Limited Optical DPSK Receiver,"* IEEE TRANSACTIONS ON COMMUNICATIONS, **Vol.37, No.1,** *January, 1989, pp.46-51*

[18] A.A. Shaar, P.A.Davies, *"Prime Sequences: Quasi-optimal Sequences For OR Channel Code Division Multiplexing"*, ELECTR ONIC
pp.888-890

[19] R.F. Serfozo, *"Conditional Poisson Processes,"* J. APPL. PROB., 9, *Sept, 1972, pp.288-302*

[20] J. Grandell, *"On Stochastic Processes Generated by a Stochastic Intensity Function"*, SKNDINAVISK AKTUARIETIDSKRIFT, *1971, pp.202-204*

[21] D.P. Brady, S. Verdú, *"Error Probability of CDMA in the Poisson Channel,"* THE 1989 IEEE/CAM INFORMATION THEORY WORKSHOP PROCEEDINGS, Cornell University, *Sec.4, p.3*